10088076 . 071902

530 Rec'd PCT/PTO 19 JUL 2002

| FORM PTO-1390 (Modified) (REV 5-93) U S DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE<br>TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371 | | ATTORNEY'S DOCKET NUMBER<br>078883-0146 |
|---|---|---|
| | | U S APPLICATION NO (If known, see 37 C F R 1 5)<br>10/088,076 |
| INTERNATIONAL APPLICATION NO.<br>PCT/GB00/03837 | INTERNATIONAL FILING DATE<br>10/05/2000 | PRIORITY DATE CLAIMED<br>10/05/1999 |

TITLE OF INVENTION
PRODUCER CELL FOR THE PRODUCTION OF RETROVIRAL VECTORS

APPLICANT(S) FOR DO/EO/US
Jason SLINGSBY , Susan Mary KINGSMAN, Jonathan ROHLL, and Andrew SLADEI.

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☐ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.
2. ☒ This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.
3. ☐ This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. ☐ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. ☐ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   - ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   - ☐ has been transmitted by the International Bureau.
   - ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)
6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2))..
7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   - ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   - ☐ have been transmitted by the International Bureau.
   - ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   - ☐ have not been made and will not be made.
8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).
11. ☐ Applicant claims small entity status under 37 CFR 1.27 .

Items 12. to 17. below concern other document(s) or information included:

12. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
13. ☒ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
14. ☐ A FIRST preliminary amendment.
    ☒ A SECOND or SUBSEQUENT preliminary amendment.
15. ☐ A substitute specification.
16. ☐ A change of power of attorney and/or address letter.
17. ☒ Other items or information: Preliminary Amendment; Statement to Support Filing and Submission; Sequence Listing (54 pgs.); and Computer Readable format

10088076 .071902

| U.S. APPLICATION NO (If known, see 37 C F R 1 50) | INTERNATIONAL APPLICATION NO | ATTORNEY'S DOCKET NUMBER |
|---|---|---|
| 10/088,076 | PCT/GB00/03837 | 078883-0146 |

| 18. ☒ The following fees are submitted: | CALCULATIONS | PTO USE ONLY |
|---|---|---|
| Basic National Fee (37 CFR 1.492(a)(1)-(5): Search Report has been prepared by the EPO or JPO ................................ $890.00 | | |
| International preliminary examination fee paid to USPTO (37 CFR 1.482) ................................ $710.00 | | |
| No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2) ................ $740.00 | | |
| Neither international preliminary examination fee (37 CFR 1.482) nor International search fee (37 CFR 1.445(a)(2)) paid to USPTO .................. $1,040.00 | | |
| International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) ............................ $100.00 | | |
| ENTER APPROPRIATE BASIC FEE AMOUNT = | | |
| Surcharge of $130.00 for furnishing the oath or declaration later than 20 Months from the earliest claimed priority date (37 CFR 1.492(e)) | $130.00 | |

| Claims | Number Filed | | Included in Basic Fee | | Extra Claims | | Rate | | |
|---|---|---|---|---|---|---|---|---|---|
| Total Claims | | - | 20 | = | 0 | × | $18.00 | $0.00 | |
| Independent Claims | | - | 3 | = | 0 | × | $84.00 | $0.00 | |
| Multiple dependent claim(s) (if applicable) | | | | | | | $280.00 | | |

| | | |
|---|---|---|
| TOTAL OF ABOVE CALCULATIONS = | $130.00 | |
| Reduction by ½ for filing by small entity, if applicable. | $0.00 | |
| SUBTOTAL = | $130.00 | |
| Processing fee of $130.00 for furnishing English translation later the 20 months from the earliest claimed priority date (37 CFR 1.492(f). + | | |
| TOTAL NATIONAL FEE = | $130.00 | |
| Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). $40.00 per property + | | |
| TOTAL FEES ENCLOSED = | $130.00 | |

07/23/2002 GFREY1 00000173 10088076

01 FC:154 130.00 OP

| | Amount to be: | |
|---|---|---|
| | refunded | $ |
| | charged | $ |

a. ☒ A check in the amount of $130.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. 19-0741 in the amount of $130.00 to the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 19-0741. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Foley & Lardner
Customer Number: 22428

22428
PATENT TRADEMARK OFFICE

Michele M Simkin
SIGNATURE

NAME MICHELE M. SIMKIN

REGISTRATION NUMBER 34,717

10088076 .072902
Rec'd PCT/PTO 20 MAR 2002

| | | |
|---|---|---|
| FORM PTO-1390 (Modified) (REV 5-93) U S DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE | | ATTORNEY'S DOCKET NUMBER |
| TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371 | | 078883-0146 10/088076 |
| | | U S APPLICATION No. Unassigned |
| INTERNATIONAL APPLICATION NO. PCT/GB00/03837 | INTERNATIONAL FILING DATE 5 October 2000 | PRIORITY DATE CLAIMED 5 October 1999 |

TITLE OF INVENTION
PRODUCER CELL FOR THE PRODUCTION OF RETROVIRAL VECTORS

APPLICANT(S) FOR DO/EO/US
Jason SLINGSBY, Susan Mary KINGSMAN, Jonathan ROHILL, and Andrew SLADE

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a FIRST submission of items concerning a filing under 35 U.S.C. 371.

2. ☐ This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.

3. ☐ This express request to begin national examination procedures (35 U.S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).

4. ☒ A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.

5. ☒ A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   - ☐ is transmitted herewith (required only if not transmitted by the International Bureau).
   - ☒ has been transmitted by the International Bureau.
   - ☐ is not required, as the application was filed in the United States Receiving Office (RO/US)

6. ☐ A translation of the International Application into English (35 U.S.C. 371(c)(2)).

7. ☒ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   - ☐ are transmitted herewith (required only if not transmitted by the International Bureau).
   - ☐ have been transmitted by the International Bureau.
   - ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   - ☒ have not been made and will not be made.

8. ☐ A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).

9. ☐ An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).

10. ☐ A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

11. ☐ Applicant claims small entity status under 37 CFR 1.27 .

Items 12. to 17. below concern other document(s) or information included:

12. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.

13. ☐ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.

14. ☒ A FIRST preliminary amendment.
    ☐ A SECOND or SUBSEQUENT preliminary amendment.

15. ☐ A substitute specification.

16. ☐ A change of power of attorney and/or address letter.

17. ☒ Other items or information: Application Data Sheet (4 pages)

10088076 .071902

JC13 Rec'd PCT/PTO 20 MAR 2002

| U.S. APPLICATION NO (If known, see 37 CFR 1.50) Unassigned 10/088076 | INTERNATIONAL APPLICATION NO PCT/GB00/03837 | ATTORNEY'S DOCKET NUMBER 078883-0146 |
|---|---|---|

| 18. ☒ The following fees are submitted: | CALCULATIONS | PTO USE ONLY |
|---|---|---|
| Basic National Fee (37 CFR 1.492(a)(1)-(5):<br>Search Report has been prepared by the EPO or JPO................................$890.00 | | |
| International preliminary examination fee paid to USPTO<br>(37 CFR 1.482)........................................................................................$710.00 | | |
| No international preliminary examination fee paid to USPTO (37 CFR 1.482)<br>but international search fee paid to USPTO (37 CFR 1.445(a)(2) ................$740.00 | | |
| Neither international preliminary examination fee (37 CFR 1.482) nor<br>International search fee (37 CFR 1.445(a)(2)) paid to USPTO ................. $1,040.00 | | |
| International preliminary examination fee paid to USPTO (37 CFR 1.482)<br>and all claims satisfied provisions of PCT Article 33(2)-(4) ...........................$100.00 | | |
| ENTER APPROPRIATE BASIC FEE AMOUNT = | $890.00 | |
| Surcharge of $130.00 for furnishing the oath or declaration later than 20<br>Months from the earliest claimed priority date (37 CFR 1.492(e)) | | |

| Claims | Number Filed | | Included in Basic Fee | | Extra Claims | | Rate | | |
|---|---|---|---|---|---|---|---|---|---|
| Total Claims | 48 | - | 20 | = | 28 | × | $18.00 | $504.00 | |
| Independent Claims | 4 | - | 3 | = | 1 | × | $84.00 | $84.00 | |
| Multiple dependent claim(s) (if applicable) | | | | | | | $280.00 | | |

| | | |
|---|---|---|
| TOTAL OF ABOVE CALCULATIONS = | $1478.00 | |
| Reduction by ½ for filing by small entity, if applicable. | $0.00 | |
| SUBTOTAL = | $1478.00 | |
| Processing fee of $130.00 for furnishing English translation later the 20 months from the earliest claimed priority date (37 CFR 1.492(f). + | | |
| TOTAL NATIONAL FEE = | $1478.00 | |
| Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). $40.00 per property + | | |
| TOTAL FEES ENCLOSED = | $1478.00 | |
| Amount to be: refunded | $ | |
| charged | $ | |

a. ☒ A check in the amount of $1478.00 to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. 19-0741 in the amount of $0.00 to the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 19-0741. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Foley & Lardner
Customer Number: 22428

22428
PATENT TRADEMARK OFFICE

Michael M Simkin
Reg No. 34,717
SIGNATURE for

NAME BERNHARD D. SAXE

REGISTRATION NUMBER 28,665

## IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

**Atty. Docket No: 078883-0146**

*In re* patent application of

SLINGSBY, JASON *et al.*

Serial No. 10/088,076

Filed: March 20, 2002

For: PRODUCER CELL FOR THE PRODUCTION OF RETROVIRAL VECTORS

### STATEMENT TO SUPPORT FILING AND SUBMISSION IN ACCORDANCE WITH 37 C.F.R. §§ 1.821-1.825

Assistant Commissioner for Patents
Washington, D.C. 20231
**Box SEQUENCE**

Sir:

In connection with a Sequence Listing submitted concurrently herewith, the undersigned hereby states that:

1. the submission, filed herewith in accordance with 37 C.F.R. § 1.821(g), does not include new matter;

2. the content of the attached paper copy and the attached computer readable copy of the Sequence Listing, submitted in accordance with 37 C.F.R. § 1.821(c) and (e), respectively, are the same; and

3. all statements made herein of their own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United

Serial No. 10/088,076

States Code and that such willful false statements may jeopardize the validity of the application or any patent resulting therefrom.

Respectfully submitted,

July 16, 2002
Date

James A. Coburn

**HARBOR CONSULTING**
Intellectual Property Services
1500A Lafayette Road
Suite 262
Portsmouth, N.H.
800-318-3021

PATENT
ATTORNEY DOCKET NO.: 078883/0146

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | |
|---|---|---|
| In re Application of: | ) | |
| Slingsby et al. | ) | |
| Appln. No.: 10/088,076 | ) | Group Art Unit: Unknown |
| Filed: March 20, 2002 | ) | Examiner: Unknown |

For: PRODUCER CELL FOR THE PRODUCTION OF RETROVIRAL VECTORS

Commissioner for Patents
Washington, D.C. 20231

Sir:

**PRELIMINARY AMENDMENT**

In response to the Notice to Comply with Requirements for Applications Containing Sequence Disclosures, mailed on May 20, 2002, please amend this application as follows.

**IN THE SPECIFICATION:**

Please replace the following paragraph with the amended version that follows. In accordance with 37 C.F.R. § 1.21(c)(2), a marked-up version of the specification is set forth at the end of this response under the heading "Marked-up Version of the Specification."

Please delete the paragraph at page 19, lines 17-21, and replace it with the following paragraph:

Figure 16 shows the alignment of leader and gag regions present in vectors pONY4Z (SEQ ID NO: 63), 8Z (SEQ ID NO: 64) and ATG mutated 8Z (SEQ ID NO: 65) vector. The latter is referred to as pONY8ZA. The sequences aligned are from the NarI site in the leader to the Xba site between the EIAV gag sequence and the CMV promoter. Sequences in the leader are shown in italic and a space is present upstream of the position of the gag ATG; and

ATTORNEY DOCKET NO.: 078883/0146
APPLN. NO.: 10/088,076

## REMARKS

Applicants respectfully request formal examination of this application.

Applicants have amended the specification to identify amino acid sequences according to their SEQ ID NOs. As the amendments do not introduce new matter, entry thereof by the Examiner is respectfully requested.

Favorable reconsideration of the application as amended is respectfully requested. The Examiner is invited to contact the undersigned by telephone if it is felt that a telephone interview would advance the prosecution of the present application.

If there are any fees due in connection with the filing of this Preliminary Amendment, please charge the fees to our Deposit Account No. 19-0741. If a fee is required for an extension of time under 37 C.F.R. § 1.136 not accounted for above, such an extension is requested and the fee should also be charged to our Deposit Account.

Respectfully submitted,

By Michele M. Simkin

Michele M. Simkin
Attorney for Applicant
Registration No. 34,717

Date July 19, 2002

FOLEY & LARDNER
Washington Harbour
3000 K Street, N.W., Suite 500
Washington, D.C. 20007-5109
Telephone: (202) 672-5538
Facsimile: (202) 672-5399

ATTORNEY DOCKET NO.: 078883/0146
APPLN. NO.: 10/088,076

**MARKED-UP VERSION OF THE SPECIFICATION**

Please delete the paragraph at page 19, lines 17-21, and replace it with the following paragraph:

Figure 16 shows the alignment of leader and gag regions present in vectors pONY4Z **(SEQ ID NO: 63)**, 8Z **(SEQ ID NO: 64)** and ATG mutated 8Z **(SEQ ID NO: 65)** vector. The latter is referred to as pONY8ZA. The sequences aligned are from the NarI site in the leader to the Xba site between the EIAV gag sequence and the CMV promoter. Sequences in the leader are shown in italic and a space is present upstream of the position of the gag ATG; and

10/088076
JC13 Rec'd PCT/PTO 20 MAR 2002

Atty. Dkt. No. 078883-0146

*IN THE UNITED STATES PATENT AND TRADEMARK OFFICE*

Applicants: Jason SLINGSBY et al.

Title: PRODUCER CELL FOR THE PRODUCTION OF RETROVIRAL VECTORS

Appl. No.: Unassigned

Filing Date: 03/20/2002

Examiner: Unassigned

Art Unit: Unassigned

**PRELIMINARY AMENDMENT**

Commissioner for Patents
Box PCT
Washington, D.C. 20231

Sir:

Prior to examination of the above-identified application, Applicants respectfully request that the above-identified application be amended as follows:

**In the Claims:**

Please cancel claims 42 and 50.

In accordance with 37 C.F.R. §1.21, please substitute for claims 3, 4, 6 – 8, 10 – 13, 15, 16, 22, 24, 26, 27, 30, 34, 38, 40, 41, 43, and 47 – 49, the following rewritten version of the same claims. The changes are shown explicitly in the attached "Marked-up Version of Amended Claims."

3. (Once Amended) A method according to claim 1, wherein the construct further comprises the 5'LTR and/or the packaging signal.

4. (Once Amended) A method according to claim 1, wherein the construct LTR is a heterologous regulatable LTR.

6. (Once Amended) A method according to claim 1, wherein the construct LTR is inactive.

7. (Once Amended) A method according to claim 1, wherein the provirus comprises an NOI encoding a selectable marker, which NOI is flanked by recombianse recognition sites.

8. (Once Amended) A method according to claim 1, wherein the provirus comprises an internal 5′ LTR upstream of the recombinase site or the 5′ recombinase site where there is more than one site.

10. (Once Amended) A method according to claim 1, wherein the U3 region of the 5′ LTR and/or the U3 region of the second internal 5′ LTR comprises a heterologous promoter.

11. (Once Amended) A method according to claim 1, wherein the provirus comprises two recombinase recognition sites and as a preliminary step, the recombinase is expressed in a host cell such that the nucleotide sequence present between the two sites is excised.

12. (Once Amended) A method according to claim 1, wherein the producer cell is a high titre producer cell, capable of producing at least $10^6$ retrovirus particles per ml.

13. (Once Amended) A method according to claim 1, wherein the provirus is a lentivirus.

15. (Once Amended) A method according to claim 2, wherein the provirus further comprises a second NOI.

16. (Once Amended) A producer cell obtainable by the method of claim 1.

22. (Once Amended) A producer cell according to claim 18, wherein the third LTR is transcriptionally quiscent.

24. (Once Amended) A producer cell according to claim 20, wherein the first NOI is a selectable marker.

26. (Once Amended) A producer cell according to claim 25, wherein the second LTR comprises a deletion in the U3 sequences in the 3′ LTR.

27. (Once Amended) A producer cell according to claim 25, wherein the second NOI comprises a coding sequence operably linked to a promotor.

30. (Once Amended) A method for producing a high titre regulatable retroviral vector, the method comprising:

(i) providing a derived producer cell comprising integrated into its genome a first vector;

(ii) introducing a second vector into the derived producer cell using a recombinase assisted method;

wherein the derived producer cell comprises a retroviral vector comprising in the 5′ to 3′ direction a first LTR (5′ LTR); a second NOI operably linked to a second LTR (regulatable 3′ LTR); and a third LTR (3′ LTR); wherein the third LTR is positioned downstream of the second LTR in the derived producer cell.

34. (Once Amended) A process for preparing a regulated retroviral vector, comprising performing the method according to claim 30 and preparing a quantity of the regulated retroviral vector.

38. (Once Amended) A regulated retroviral vector according to claim 36, wherein the target site is a cell.

40. (Once Amended) A regulated retroviral vector according to claim 35, in combination with a pharmaceutically acceptable carrier.

41. (Once Amended) A medicament for diagnostic and/or therapeutic and/or medical applications, comprising a regulated retroviral vector according to claim 35.

43. (Once Amended) A derived stable producer cell capable of expressing regulated retroviral vectors according to claim 35.

47. (Once Amended) A nucleic acid vector according to claim 45, further comprising a 5′ LTR and/or a packaging signal.

48. (Once Amended) A nucleic acid vector according to claim 45, wherein the LTR is a heterologous regulatable LTR.

49. (Once Amended) A nucleic acid vector according to claim 45, wherein the LTR is transcriptionally quiscent.

Atty. Dkt. No. 078883-0146

## REMARKS

Applicants respectfully request that the foregoing amendments be made prior to examination of the present application. The amendments are made to correct multiple dependencies and do not change the scope of the invention.

Respectfully submitted,

Michael M Dsela

By Reg No. 34,717 /BS

Bernhard D. Saxe
Attorney for Applicants
Registration No. 28,665

Date March 20, 2002

FOLEY & LARDNER
Customer Number: 22428

22428

PATENT TRADEMARK OFFICE
Telephone: (202) 672-5427
Facsimile: (202) 672-5399

**MARKED UP VERSION OF AMENDED CLAIMS**

3. (Once Amended) A method according to claim 1, [or claim 2] wherein the construct further comprises the 5'LTR and/or the packaging signal.

4. (Once Amended) A method according to [any one of claims] claim 1 [to 3], wherein the construct LTR is a heterologous regulatable LTR.

6. (Once Amended) A method according to [any one of claims] claim 1, [to 3] wherein the construct LTR is inactive.

7. (Once Amended) A method according to [any one of the preceding claims] claim 1, wherein the provirus comprises an NOI encoding a selectable marker, which NOI is flanked by recombianse recognition sites.

8. (Once Amended) A method according to [any one of the precedings claims] claim 1, wherein the provirus comprises an internal 5' LTR upstream of the recombinase site or the 5' recombinase site where there is more than one site.

10. (Once Amended) A method according to [any one of the preceding claims] claim 1, wherein the U3 region of the 5' LTR and/or the U3 region of the second internal 5' LTR comprises a heterologous promoter.

11. (Once Amended) A method according to [any one of the preceding claims] claim 1, wherein the provirus comprises two recombinase recognition sites and as a preliminary step, the recombinase is expressed in a host cell such that the nucleotide sequence present between the two sites is excised.

12. (Once Amended) A method according to [any one of the preceding claims] claim 1, wherein the producer cell is a high titre producer cell, capable of producing at least $10^6$ retrovirus particles per ml.

13. (Once Amended) A method according to [any one of the preceding claims] claim 1, wherein the provirus is a lentivirus.

15. (Once Amended) A method according to [any one of claims] claim 2 [- 14], wherein the provirus further comprises a second NOI.

16. (Once Amended) A producer cell obtainable by the method of [any one of claims] claim 1 [to 15].

22. (Once Amended) A producer cell according to [any one of claims] claim 18, [to 21] wherein the third LTR is transcriptionally quiscent.

24. (Once Amended) A producer cell according to [any one of claims] claim 20, [to 23] wherein the first NOI is a selectable marker.

26. (Once Amended) A producer cell according to claim 25, wherein the second LTR comprises a deletion in the U3 sequences in the 3′ LTR.

27. (Once Amended) A producer cell according to claim 25 [or claim 26], wherein the second NOI comprises a coding sequence operably linked to a promotor.

30. (Once Amended) A method for producing a high titre regulatable retroviral vector, the method comprising [the steps of]:

(i) providing a derived producer cell comprising integrated into its genome a first vector;

(ii) introducing a second vector into the derived producer cell using a recombinase assisted method;

wherein the derived producer cell comprises a retroviral vector comprising in the 5′ to 3′ direction a first LTR (5′ LTR); a second NOI operably linked to a second LTR (regulatable

3′ LTR); and a third LTR (3′ LTR); wherein the third LTR is positioned downstream of the second LTR in the derived producer cell.

34. (Once Amended) A process for preparing a regulated retroviral vector, [as defined in claim 17] comprising performing the method according to [any one of claims] claim 30 [to 33] and preparing a quantity of the regulated retroviral vector.

38. (Once Amended) A regulated retroviral vector according to claim 36, [or claim 37] wherein the target site is a cell.

40. (Once Amended) [Use of a] A regulated retroviral vector according to [any one of claims] claim 35, [to 38] in [the manufacture of a pharmaceutical composition to deliver an NOI to a target site] combination with a pharmaceutically acceptable carrier.

41. (Once Amended) [Use of a regulated retroviral vector according to any one of claims 35 to 38 in the manufacture of a] A medicament for diagnostic and/or therapeutic and/or medical applications , comprising a regulated retroviral vector according to claim 35.

43. (Once Amended) A derived stable producer cell capable of expressing regulated retroviral vectors according to claim[s] 35 to [38].

47. (Once Amended) A nucleic acid vector according to claim 45, [or claim 46] further comprising a 5′ LTR and/or a packaging signal.

48. (Once Amended) A nucleic acid vector according to [any one of claims] claim 45, [to 47] wherein the LTR is a heterologous regulatable LTR.

49. (Once Amended) A nucleic acid vector according to [any one of claims] claim 45, [to 47] wherein the LTR is transcriptionally quiscent.

PCT/PTO 20 MAR 2002



10/088076

# PRODUCER CELL FOR THE PRODUCTION OF RETROVIRAL VECTORS

## FIELD OF THE INVENTION

The present invention relates to retroviral vectors, in particular to high titre regulatable retroviral vectors.

## BACKGROUND TO THE INVENTION

Retroviruses have been proposed as a delivery system (otherwise expressed as a delivery vehicle or delivery vector) for *inter alia* the transfer of a nucleotide sequence of interest (NOI), or a plurality of NOIs, to one or more sites of interest. The transfer can occur *in vitro, ex vivo, in vivo,* or combinations thereof. When used in this fashion, the retroviruses are typically called retroviral vectors or recombinant retroviral vectors. Retroviral vectors have been exploited to study various aspects of the retrovirus life cycle, including receptor usage, reverse transcription and RNA packaging (reviewed by Miller, 1992 Curr Top Microbiol Immunol 158:1-24).

In a typical recombinant retroviral vector for use in gene therapy, at least part of one or more of the *gag, pol* and *env* protein coding regions may be removed from the virus. This makes the retroviral vector replication-defective. The removed portions may even be replaced by a NOI in order to generate a virus capable of integrating its genome into a host genome but wherein the modified viral genome is unable to propagate itself due to a lack of structural proteins. When integrated in the host genome, expression of the NOI occurs - resulting in, for example, a therapeutic effect. Thus, the transfer of a NOI into a site of interest is typically achieved by: integrating the NOI into the recombinant viral vector; packaging the modified viral vector into a virion coat; and allowing transduction of a site of interest - such as a targetted cell or a targetted cell population.

It is possible to propagate and isolate quantities of retroviral vectors (e.g. to prepare suitable titres of the retroviral vector) for subsequent transduction of, for example, a site of interest by using a combination of a packaging or helper cell line and a recombinant vector.

In some instances, propagation and isolation may entail isolation of the retroviral *gag, pol* and *env* genes and their separate introduction into a host cell to produce a "packaging cell line". The packaging cell line produces the proteins required for packaging retroviral RNA but it does not produce RNA-containing retroviral vectors. However, when a recombinant vector carrying a NOI and a *psi* region is introduced into the packaging cell line, the helper proteins can package the *psi*-positive recombinant vector to produce the recombinant virus stock. This can be used to infect cells to introduce the NOI into the genome of the cells. The recombinant virus whose genome lacks all genes required to make viral proteins can infect only once and cannot propagate. Hence, the NOI is introduced into the host cell genome without the generation of potentially harmful retrovirus. A summary of the available packaging lines is presented in "Retroviruses" (1997 Cold Spring Harbour Laboratory Press Eds: JM Coffin, SM Hughes, HE Varmus pp 449).

Retroviral packaging cell lines have been developed to produce retroviral vectors. These cell lines are designed to express three components, which may be located on three separate expression constructs. The *gag/pol* expression construct encodes structural and enzymatic components required in particle formation, maturation, reverse transcription and integration. The envelope (*env*) construct expresses a retroviral or non-retroviral envelope protein, which mediates viral entry into cells by binding to its cognate receptor. The third expression construct produces the retroviral RNA genome containing a *psi* region, which is packaged into mature, enveloped retroviral particles.

It has been observed that different methods, such as electroporation, transfection and retroviral transduction, which have been used to introduce the retroviral expression construct for the RNA genome, termed "the genome", into packaging cells produce different results. These different results can include different end points or "yield" of retroviral producer lines resulting from the derived cell lines. Moreover, electroporation and transfection methods can be problematic in the sense that the titre levels are not always at a satisfactory level.

By way of example, the transfection of a plasmid DNA construct into packaging cells from a MLV packaging cell line of human origin, called FLYA13, yielded low retroviral

vector titres even when different transfection reagents such as calcium phosphate precipitation and fugene transfection reagent were used. The average titres from selected stably transfected cell lines clones ranged from about $10^3$ to about $10^4$ per ml. In addition, clones generated by electroporation of constructs gave similar titres of from about $10^3$ to about $10^4$ per ml with no clones identified producing at $>10^5$ per ml. However, when MLV vector particles are prepared in a transient transfection system with a different envelope pseudotype to the packaging cell, and used to transduce a retroviral packaging cell, stably transduced cell lines made by this transduction method produce retrovirus at $10^6$ to $10^7$ per ml. Therefore, these results suggest that retroviral transduction is a preferred method for genome introduction into packaging cell lines in order to generate high titre producer cell lines. However, when retroviral transduction is used to introduce a regulated/inactivated retroviral vector genome into packaging cell lines, the regulated retroviral vectors may not be produced in sufficient quantities from these cell lines.

By way of example, some retroviral vectors may comprise (i) internal expression constructs which are themselves regulated or (ii) regulated elements which are present in retroviral 3' LTR sequences, either by design or by their nature. Examples of these regulated vectors include but are not limited to hypoxic regulated vectors and self inactivating (SIN) vectors. If transduced producer cell lines are generated with these regulated vectors, the regulated or inactivated 3' U3 sequence of the LTR is copied to the 5' LTR by the process of retroviral reverse transcription and integration. Therefore, in the producer cell line, the 5' U3 promoter sequence directing expression of retroviral RNA genomes is identical to the regulated or inactivated 3' U3 promoter. This will result in very low levels of retroviral genome production and consequently low titres of functional retrovirus vectors being produced.

One example of such a regulated retroviral system includes MLV and lentivirus vector constructs where the 3' retroviral U3 enhancer element is replaced with a hypoxia responsive element (HRE) or other physiologically regulated, tumour specific or tissue-specific promoters. When these vectors are used to make a transduced producer cell line, the 3' U3 sequence containing the HRE element is copied to the 5' LTR position and retroviral genomes will only be produced under hypoxic conditions or chemical mimics

of hypoxia, such as heavy metal ions and desferrioxamine. Such a requirement for "induction for retroviral production" is not preferable as the different hypoxia induction protocols negatively affect retroviral producer cell viability.

By way of further example, lentivector U3 enhancers are dependent on the transactivator TAT for transcriptional activation. Therefore, a lentivector producer cell line generated by transduction requires the presence of TAT for high level expression of the lentivector genome construct. The expression of TAT is not preferable in such a packaging cell line and therefore, in the absence of TAT, only very low titres will be produced from transduced producer cells generated in this way.

Another example of a regulated retroviral systems includes MLV or lentivirus self-inactivating (SIN) vectors. These vectors contain deletions of the elements in their 3' U3 sequences responsible for transcriptional activity. Therefore, on transduction of target cells, the transcriptionally inactive 3' U3 sequence is copied to the 5' LTR position. In standard configurations, an internal expression cassette directs therapeutic or marker gene expression. However, if SIN vectors are used to make a transduced retroviral producer line, there will be no transcriptional elements present to direct high levels of retroviral RNA genome expression.

Although it is possible to carry out retroviral transduction with much lower-titre vector stocks, for practical reasons, high-titre retrovirus is desirable, especially when a large number of cells must be infected. In addition, high titres are a requirement for transduction of a large percentage of certain cell types. For example, the frequency of human hematopoietic progenitor cell infection is strongly dependent on vector titre, and useful frequencies of infection occur only with very high-titre stocks (Hock and Miller 1986 Nature 320: 275-277; Hogge and Humphries 1987 Blood 69: 611-617). In these cases, it is not sufficient simply to expose the cells to a larger volume of virus to compensate for a low virus titre. On the contrary, in some cases, the concentration of infectious vector virions may be critical to promote efficient transduction.

## SUMMARY OF THE INVENTION

We have now shown that it is possible to obtain transduced producer cells capable of producing high titre regulated retroviral vectors by replacing at least the 3'LTR of the integrated provirus using a recombinase based system. Thus whereas with the prior art, the U3 region of the 3'LTR is the same as that of the U3 region of the 5' LTR (and vice versa for the U5 region) in the provirus due to the way in which the viral vector integrates, the introduction of, for example, a replacement 3'LTR results in a provirus that has a 5'LTR and a 3'LTR that differ. The packaged viral vectors produced by transcription of the proviral genome within the producer cells may then ultimately be used to transduce target cells where the regulatable sequences present in the 3'LTR of the provirus in the producer cells are then present in the 5'LTR of the provirus in the target cells and consequently regulate transcription from the provirus as required.

This allows the introduction of a 3'LTR, for example a regulatable 3'LTR, into the provirus that was not desirable in the original viral vector used to transduce the producer cells since the consequential appearance of the regulatable 3'LTR U3 sequences in the 5'LTR in the provirus may lead to a reduced viral titre.

Consequently, the present invention allows transduced producer cells to be constructed that are capable of producing high titre regulated retroviral vectors by virtue of comprising a 5'LTR that directs high level expression of the viral genome in the producer cell and a different 3'LTR which as a result of the process of integration into a target cell will then result in a provirus in the target cell genome that exhibits regulatable expression.

In particular, the present invention allows the modification of a provirus integrated into the genome of the producer cells that have been selected for their high titre virus production such that the resulting packaged viral particles produced from the provirus may be used to transduce target cells resulting in a provirus integrated into the genome of the target cells that has a different, and preferably regulatable 5'LTR to that of the producer cell provirus.

The present invention is not limited to replacement of the 3'LTR of the provirus in the high titre producer cells, but may also include replacement of the 5'LTR and other viral sequences and/or the introduction of NOIs by the use of suitable constructs, as shown in the Figures.

Accordingly, the present invention provides a method of modifying a producer cell which producer cell comprises integrated into its genome a provirus which provirus comprises one or more recombinase recognition sequences within or upstream of its 3' LTR, the method comprising: introducing into the cell a construct comprising a 5' recombinase recognition sequence, an LTR and a 3' recombinase recognition sequence in that order, in the presence of a recombinase which is capable of acting on the recombinase recognition site(s) such that the nucleotide sequence between the 5' and 3' recombinase recognition sequences in the construct is introduced into the provirus.

Preferably the LTR is a heterologous regulatable LTR.

The present invention further provides a nucleic acid vector comprising a 5' recombinase recognition sequence, a regulatable LTR and a 3' recombinase recognition sequence in that order.

In any of the above aspects and embodiments of the invention, preferably the construct, nucleic acid molecule and/or nucleic acid vector further comprises at least one NOI between the 5' recombinase recognition sequence and the regulatable LTR.

Preferably the construct, nucleic acid molecule and/or nucleic acid vector further comprises a 5'LTR and/or a packaging signal

In one embodiment of the invention, the LTR is inactive/transcriptionally quiescent.

The construct, nucleic acid molecule and/or nucleic acid vector of the invention may be used in a recombinase assisted method to introduce a regulated LTR into a proviral genome integrated into a producer cell genome.

The present invention also provides a producer cell obtainable by the method of the invention, preferably a high titre producer cells. Also provided is an infectious retroviral particle obtained by the above method.

The present invention further provides a high titre producer cell comprising integrated into its genome a provirus, which provirus comprises a recombinase recognition site, a 5' LTR and a 3'LTR which 3'LTR differs from the 5'LTR. Such a producer cell will typically have been produced by the method of the invention.

Preferably the 5'LTR and the 3'LTR referred to for the purposes of comparison are both "active". The term "active" within the present context means transcriptionally active, that is to say, the 5'LTR comprises a promoter that directs transcription of the viral genome and the 3'LTR comprises a transcriptional stop sequence to terminate transcription. This distinction is relevant since if a provirus produced by the method of the invention comprises more than one 5' LTR or 3'LTR, at least one but not all must be active to allow viral production. Further, if the provirus comprises more than one 3'LTR then it is generally the upstream one that will be active since transcription will tend not to read through to the downstream 3' LTR.

In addition, where the method of the invention results in an insertion of a 3'LTR upstream of the original 3'LTR, the comparison should be performed between the additional 3'LTR and the original 5'LTR and not the two original LTRs. Thus it is permitted to have a 5'LTR and 3'LTR within the same provirus that are the same provided that there is also a 5'LTR and 3'LTR that differ.

In another aspect, the present invention provides a derived producer cell comprising integrated into its genome a retroviral vector comprising in the 5' to 3' direction a first 5' LTR; a second NOI operably linked to a second regulatable 3' LTR; and a third 3'LTR; wherein the third 3'LTR is positioned downstream of the second regulatable 3'LTR in the producer cell.

Preferably the first 5' LTR comprising 5'R and 5' U5 sequences is derivable from a first vector; the second NOI operably linked to a second regulatable 3' LTR is derivable from a second vector; and the third 3'LTR is derivable from the first vector.

In a preferred embodiment, the first vector further comprises an internal LTR located upstream of the first NOI and downstream of the packaging signal wherein the internal LTR comprises a heterologous U3 sequence linked to heterologous R and U5 sequences.

Preferably the heterologous R and U5 sequences are lentiviral derivable R and U5 sequences, such as EIAV R and U5 sequences.

In a further preferred embodiment, the third 3'LTR is transcriptionally active but expression is directed away from the second regulatable 3'LTR.

In another embodiment, the second vector comprises a second NOI operably linked to a second regulatable 3'LTR comprising at least one recombinase recognition sequence. Preferably the second regulatable 3'LTR comprises a deletion in the U3 sequences in the 3'LTR.

Preferably, the second NOI comprises a discistronic construct, more preferably a discistronic construct comprising a therapeutic gene, an internal ribosomal entry site (IRES) and a reporter gene.

The present invention further provides in another embodiment, a method for producing a high titre regulatable retroviral vector, the method comprising the steps of:
(i) providing a derived producer cell comprising integrated into its genome a first vector;
(ii) introducing a second vector into the derived producer cell using a recombinase assisted method; wherein the derived producer cell comprises a retroviral vector comprising in the 5' to 3' direction a first 5' LTR; a second NOI operably linked to a second regulatable 3' LTR; and a third 3'LTR; wherein the third 3'LTR is positioned downstream of the second regulatable 3'LTR in the derived producer cell.

The present invention also provides the use of a recombinase assisted mechanism to introduce a regulated 3'LTR into a derived producer cell line to produce a high titre regulated retroviral vector.

Aspects of the present invention are also presented in the accompanying claims and in the following description and discussion.

These aspects are presented under separate section headings. However, it is to be understood that the teachings under each section heading are not necessarily limited to that particular section heading.

## DETAILED DESCRIPTION OF THE INVENTION

The present invention is advantageous because:

(i) it enables regulated retroviral vectors to be produced at high titres from transduced producer cell lines.

(ii) it removes the uncertainty associated with the process of producer cell line derivation and the necessity to screen large numbers of producer cell lines each time a new retroviral expression construct is introduced into a producer cell line.

(iii) it greatly facilitates the generation of high titre retroviral stocks without the use of marker genes (such as but not limited to β-galactosidase, green fluorescent protein) and antibiotic resistance genes.

(iv) it avoids the derivation of low titre transfected producer cell lines or the use of hypoxic conditions or chemical mimics for production from traditionally derived transduced producer lines.

(v) it enables the production of SIN vectors by stable cell line producer technology. Previously, SIN vectors have not been amenable to production by stable cell line producer technology because the deletion of the 3'U3 sequence resulted in at least a tenfold lower

titre of self-inactivating (SIN) vectors in comparison with vectors having intact LTRs. Consequently, SIN vectors have had to be prepared using transfection-based transient expression systems.

PRODUCER CELL

The high titre regulated retroviral vector particles of the present invention are typically generated in a suitable producer cell. Producer cells are generally mammalian cells but can be, for example, insect cells. A producer cell may be a packaging cell containing the virus structural genes, normally integrated into its genome into which the regulated retroviral vectors of the present invention are introduced. Alternatively the producer cell may be transfected with nucleic acid sequences encoding structural components, such as *gag/pol/env* on one or more vectors such as plasmids, adenovirus vectors, herpes viral vectors or any method known to deliver functional DNA into target cells. The vectors according to the present invention are then introduced into the packaging cell by the methods of the present invention.

As used herein, the term "producer cell" or "vector producing cell" refers to a cell which contains all the elements necessary for production of regulated retroviral vector particles and regulated retroviral delivery systems.

Preferably, the producer cell is obtainable from a stable producer cell line.

Preferably, the producer cell is obtainable from a derived stable producer cell line.

Preferably, the producer cell is obtainable from a derived producer cell line

As used herein, the term "derived producer cell line" is a transduced producer cell line which has been screened and selected for high expression of a marker gene. Such cell lines contain retroviral insertions in integration sites that support high level expression from the retroviral genome. The term "derived producer cell line" is used interchangeably with the term "derived stable producer cell line" and the term "stable producer cell line

Preferably the derived producer cell line includes but is not limited to a retroviral and/or a lentiviral producer cell.

Preferably the derived producer cell line is an HIV or EIAV producer cell line, more preferably an EIAV producer cell line.

Preferably the envelope protein sequences, and nucleocapsid sequences are all stably integrated in the producer and/or packaging cell. However, one or more of these sequences could also exist in episomal form and gene expression could occur from the episome.

PACKAGING CELL

As used herein, the term "packaging cell" refers to a cell which contains those elements necessary for production of infectious recombinant virus which are lacking in a recombinant viral vector. Typically, such packaging cells contain one or more expression cassettes which are capable of expressing viral structural proteins (such as *gag, pol* and *env*) but they do not contain a packaging signal.

The term "packaging signal" which is refered to interchangeably as "packaging sequence" or "*psi*" is used in reference to the non-coding sequence required for encapsidation of retroviral RNA strands during viral particle formation.

Packaging cell lines suitable for use with the above-described vector constructs may be readily prepared (see also WO 92/05266), and utilised to create producer cell lines for the production of retroviral vector particles. As already mentioned, a summary of the available packaging lines is presented in "Retroviruses" (1997 Cold Spring Harbour Laboratory Press Eds: JM Coffin, SM Hughes, HE Varmus pp 449).

The packaging cell lines are useful for providing the gene products necessary to encapsidate and provide a membrane protein for a high titre regulated retrovirus vector and regulated nucleic gene delivery vehicle production. When regulated retrovirus sequences are introduced into the packaging cell lines, such sequences are encapsidated

with the nucleocapsid (*gag/pol*) proteins and these units then bud through the cell membrane to become surrounded in cell membrane and to contain the envelope protein produced in the packaging cell line. These infectious regulated retroviruses are useful as infectious units *per se* or as gene delivery vectors.

The packaging cell may be a cell cultured *in vitro* such as a tissue culture cell line. Suitable cell lines include but are not limited to mammalian cells such as murine fibroblast derived cell lines or human cell lines. Preferably the packaging cell line is a human cell line, such as for example: HEK293, 293-T, TE671, HT1080.

Alternatively, the packaging cell may be a cell derived from the individual to be treated such as a monocyte, macrophage, blood cell or fibroblast. The cell may be isolated from an individual and the packaging and vector components administered *ex vivo* followed by re-administration of the autologous packaging cells.

Methods for introducing retroviral packaging and vector components into packaging/producer cells are described in the present invention.

Preferably the method of the present invention utilises a recombinase assisted mechanism.

Preferably the method of the present invention utilises a recombinase assisted mechanism which facilitates the production of high titre regulated retroviral vectors from the producer cells of the present invention.

RECOMBINASE ASSISTED MECHANISM

As used herein, the term "recombinase assisted system" includes but is not limited to a system using the Cre recombinase / loxP recognition sites of bacteriophage P1 or the site-specific FLP recombinase of *S. cerevisiae* which catalyses recombination events between 34 bp FLP recognition targets (FRTs).

The site-specific FLP recombinase of *S. cerevisiae* which catalyses recombination events between 34 bp FLP recognition targets (FRTs) has been configured into DNA constructs

in order to generate high level producer cell lines using recombinase-assisted recombination events (Karreman *et al*. (1996) NAR 24, 1616-1624). A similar system has been developed using the Cre recombinase / loxP recognition sites of bacteriophage P1. This was configured into a retroviral genome such that high titre retroviral producer cell lines were generated (Vanin *et al.* (1997) J Virol 71, 7820-7826). However, the use of the second method (Vanin *et al ibid*) has centered around the exchange of the central portions of a retroviral cassette using a recombinase-assisted system. Moreover, these methods have used genes encoding selectable markers such as $neo^R$ and $puro^R$ (Vanin *et al ibid*) and luciferase and $puro^R$ linked by an IRES sequence (Karreman *et al ibid*). Karreman and Vanin do not demonstrate or *suggest* that: (i) a regulated or inactive 3'U3 sequence of the 3'LTR can be introduced into a producer cell via a recombinase-assisted mechanism or (ii) that therapeutic genes under the control of a regulated LTR may be introduced into a producer cell line via a recombinase assisted step. Vanin *et al ibid* suggests that his Cre-mediated recombination approach to retroviral producer cell line production may be used in combination with other modifications which should result in improved vector performance. Vanin *et al ibid* also suggests that his approach provides a means to generate high titre SIN vectors. However, there is no worked example and in fact no enabling disclosure because the skilled person would not have been aware, on the basis of the Vanin *et al* paper, of the necessary modifications to make the suggested approach work. Vanin *et al* makes no reference to hypoxic regulated vectors and/or regulated/inactivated lentiviral vectors.

LTRs

As already indicated, each retroviral genome comprises genes called *gag, pol* and *env* which code for virion proteins and enzymes. In the provirus, these genes are flanked at both ends by regions called long terminal repeats (LTRs). The LTRs are responsible for proviral integration, and transcription. They also serve as enhancer-promoter sequences. In other words, the LTRs can control the expression of the viral gene. Encapsidation of the retroviral RNAs occurs by virtue of a *psi* sequence located at the 5' end of the viral genome.

As used herein, the term "long terminal repeat (LTR) is used in reference to domains of base pairs located at the end of retroviral DNAs.

The LTRs themselves are identical sequences that can be divided into three elements, which are called U3, R and U5. U3 is derived from the sequence unique to the 3' end of the RNA. R is derived from a sequence repeated at both ends of the RNA and U5 is derived from the sequence unique to the 5' end of the RNA. The sizes of the three elements can vary considerably among different retroviruses.

For ease of understanding, a simple, generic structures (not to scale) of the RNA and the DNA forms of the MLV retroviral genome is presented in Figure 7 in which the elementary features of the LTRs and the relative positioning of *gag/pol* and *env* are indicated. Please note that (i) gag/pol and env are normally not spaced apart; and (ii) the overlap normally present between the *pol* and *env* genes and the poly A tail normally present at the 3' end of the RNA transcript are not illustrated in Figure 7.

As shown in Figure 7, the basic molecular organisation of an infectious retroviral RNA genome is (5') R - U5 - *gag/pol*, *env* - U3-R (3'). In a defective retroviral vector genome *gag*, *pol* and *env* may be absent or not functional. The R regions at both ends of the RNA are repeated sequences. U5 and U3 represent unique sequences at the 5' and 3' ends of the RNA genome respectively.

Upon cellular transduction, reverse transcription of the virion RNA into double stranded DNA takes place in the cytoplasm and involves two jumps of the reverse transcriptase from the 5' terminus to the 3' terminus of the template molecule. The result of these jumps is a duplication of sequences located at the 5' and 3' ends of the virion RNA. These sequences then occur fused in tandem on both ends of the viral DNA, forming the long terminal repeats (LTRs) which comprise R U5 and U3 regions. On completion of the reverse transcription, the viral DNA is translocated into the nucleus where the linear copy of the retroviral genome, called a preintegration complex (PIC), is randomly inserted into chromosomal DNA with the aid of the virion integrase to form a stable provirus. The number of possible sites of integration into the host cellular genome is very large and very widely distributed.

Preferably the retroviral genome is introduced into packaging cell lines using retroviral transduction.

Preferably retroviral vector particles (such as MLV vector particles) are prepared in a transient expression system with a different envelope pseudotype to the packaging cell, and used to transduce a retroviral packaging cell.

Preferably the retroviral transduction step identifies retroviral insertions in integration sites that support high level expression of the resulting regulated retroviral genome.

Preferably stable transduced producer cell lines made by this initial retroviral transduction step produce retrovirus at titres of at least $10^6$ per ml, such as from about $10^6$ to about $10^7$ per ml, more preferably at least about $10^7$ per ml.

HIGH TITRE

As used herein, the term “high titre” means an effective amount of a retroviral vector or particle which is capable of transducing a target site such as a cell.

As used herein, the term “effective amount” means an amount of a regulated retroviral or lentiviral vector or vector particle which is sufficient to induce expression of an NOI at a target site.

Preferably the titre is from at least $10^6$ retrovirus particles per ml, such as from about $10^6$ to about $10^7$ per ml, more preferably at least about $10^7$ retrovirus particles per ml.

TRANSCRIPTIONAL CONTROL

The control of proviral transcription remains largely with the noncoding sequences of the viral LTR. The site of transcription initiation is at the boundary between U3 and R in the left hand side LTR (as shown in Figure 7) and the site of poly (A) addition (termination) is at the boundary between R and U5 in the right hand side LTR (as shown in Figure 7). The 3’U3 sequence contains most of the transcriptional control elements of the provirus,

which include the promoter and multiple enhancer sequences responsive to cellular and in some cases, viral transcriptional activator proteins.

REGULTABLE LTRs

AN LTR present, for example, in the construct of the invention and as a 3'LTR in the provirus of the producer cell of the invention may be a native LTR or a heterologous regulatable LTR. It may also be a transcriptionally quiescent LTR for use in SIN vector technology.

As used herein, the terms "regulatable LTR" and "regulatable 3'LTR" include vectors which contain responsive elements which are present in retroviral 3' LTR sequences, either by design or by their nature. As used herein, vectors comprising a "regulatable 3'LTR" are referred to as "regulated retroviral vectors". Within the regulatable 3'LTR region, the 3'U3 sequence contains most of the transcriptional control elements of the provirus, which include the promoter and multiple enhancer sequences responsive to cellular and in some cases, viral transcriptional activator proteins.

Responsive elements include but are not limited to elements which comprise, for example, promoter and multiple enhancer sequences responsive to cellular and in some cases, viral transcriptional activator proteins and/or elements which have been modified to render them inactive. As used herein, the term "modified" includes but is not limited to silencing, disabling, mutating, deleting or removing all of the U3 sequence or a part thereof.

The term "regulated LTR" also includes an inactive LTR such that the resulting provirus in the target cell can not produce a packagable viral genome (self-inactivating (SIN) vector technology) - see the Examples and Figure 6 for a particular embodiment.

ENHANCER

As used herein, the term "enhancer" includes a DNA sequence which binds other protein components of the transcription initiation complex and thus facilitates the initiation of transcription directed by its associated promoter.

In one preferred embodiment of the present invention, the enhancer is an ischaemic like response element (ILRE).

ILRE

The term "ischaemia like response element" - otherwise written as ILRE - includes an element that is responsive to or is active under conditions of ischaemia or conditions that are like ischaemia or are caused by ischaemia. By way of example, conditions that are like ischaemia or are caused by ischaemia include hypoxia and/or low glucose concentration(s).

The term "hypoxia" means a condition under which a particular organ or tissue receives an inadequate supply of oxygen.

Ischaemia can be an insufficient supply of blood to a specific organ or tissue. A consequence of decreased blood supply is an inadequate supply of oxygen to the organ or tissue (hypoxia). Prolonged hypoxia may result in injury to the affected organ or tissue.

A preferred ILRE is an hypoxia response element (HRE).

HRE

In one preferred aspect of the present invention, there is hypoxia or ischaemia regulatable expression of the retroviral vector components. In this regard, hypoxia is a powerful regulator of gene expression in a wide range of different cell types and acts by the induction of the activity of hypoxia-inducible transcription factors such as hypoxia inducible factor-1 (HIF-1; Wang & Semenza 1993 Proc Natl Acad Sci 90:430), which

bind to cognate DNA recognition sites, the hypoxia-responsive elements (HREs) on various gene promoters. Dachs *et al* (1997 Nature Med 5: 515) have used a multimeric form of the HRE from the mouse phosphoglycerate kinase-1 (PGK-1) gene (Firth *et al* 1994 Proc Natl Acad Sci 91:6496-6500) to control expression of both marker and therapeutic genes by human fibrosarcoma cells in response to hypoxia *in vitro* and within solid tumours *in vivo* (Dachs *et al ibid*).

Hypoxia response enhancer elements (HREEs) have also been found in association with a number of genes including the erythropoietin (EPO) gene (Madan *et al* 1993 Proc Natl Acad Sci 90: 3928; Semenza and Wang 1992 Mol Cell Biol 1992 12: 5447-5454). Other HREEs have been isolated from regulatory regions of both the muscle glycolytic enzyme pyrivate kinase (PKM) gene (Takenaka *et al* 1989 J Biol Chem 264: 2363-2367), the human muscle-specific β-enolase gene (ENO3; Peshavaria and Day 1991 Biochem J 275: 427-433 ) and the endothelin-1 (ET-1) gene (Inoue *et al* 1989 J Biol Chem 264: 14954-14959).

Preferably the HRE of the present invention is selected from, for example, the erythropoietin HRE element (HREE1), muscle pyruvate kinase (PKM), HRE element, phosphoglycerate kinase (PGK) HRE, B-enolase (enolase 3; ENO3) HRE element, endothelin-1 (ET-1)HRE element and metallothionein II (MTII) HRE element.

RESPONSIVE ELEMENT

Preferably the ILRE is used in combination with a transcriptional regulatory element , such as a promoter, which transcriptional regulatory element is preferably active in one or more selected cell type(s), preferably being only active in one cell type.

As outlined above, this combination aspect of the present invention is called a responsive element.

Preferably the responsive element comprises at least the ILRE as herein defined.

Non-limiting examples of such a responsive element are presented as OBHRE1 and XiaMac. Another non-limiting example includes the ILRE in use in conjunction with an MLV promoter and/or a tissue restricted ischaemic responsive promoter. These responsive elements are disclosed in WO99/15684.

Other examples of suitable tissue restricted promoters/enhancers are those which are highly active in tumour cells such as a promoter/enhancer from a *MUC*1 gene, a *CEA* gene or a *5T4* antigen gene. The alpha fetoprotein (AFP) promoter is also a tumour-specific promoter. One preferred promoter-enhancer combination is a human cytomegalovirus (hCMV) major immediate early (MIE) promoter/enhancer combination.

PROMOTER

The term "promoter" is used in the normal sense of the art, e.g. an RNA polymerase binding site.

The promoter may be located in the retroviral 5' LTR to control the expresssion of a cDNA encoding an NOI.

Preferably the NOI is capable of being expressed from the retrovirus genome such as from endogenous retroviral promoters in the long terminal repeat (LTR)

Preferably the NOI is expressed from a heterologous promoter to which the heterologous gene or sequence is operably linked.

Alternatively, the promoter may be an internal promoter.

Preferably the NOI is expressed from an internal promoter.

Vectors containing internal promoters have also been widely used to express multiple genes. An internal promoter makes it possible to exploit promoter/enhancer combinations other than those found in the viral LTR for driving gene expression. Multiple internal promoters can be included in a retroviral vector and it has proved

possible to express at least three different cDNAs each from its own promoter (Overell *et al* 1988 Mol Cell Biol 8: 1803-1808). Internal ribosomal entry site (IRES) elements have also been used to allow translation of multiple coding regions from either a single mRNA or from fusion proteins that can then be expressed from an open reading frame.

TISSUE SPECIFIC PROMOTERS

The promoter of the present invention may be constitutively efficient, or may be tissue or temporally restricted in their activity.

Preferably the promoter is a constitutive promoter such as CMV.

Preferably the promoters of the present invention are tissue specific.

That is, they are capable of driving transcription of a NOI or NOI(s) in one tissue while remaining largely "silent" in other tissue types.

The term "tissue specific" means a promoter which is not restricted in activity to a single tissue type but which nevertheless shows selectivity in that they may be active in one group of tissues and less active or silent in another group.

The level of expression of an NOI or NOIs under the control of a particular promoter may be modulated by manipulating the promoter region. For example, different domains within a promoter region may possess different gene regulatory activities. The roles of these different regions are typically assessed using vector constructs having different variants of the promoter with specific regions deleted (that is, deletion analysis). This approach may be used to identify, for example, the smallest region capable of conferring tissue specificity or the smallest region conferring hypoxia sensitivity.

A number of tissue specific promoters, described above, may be particularly advantageous in practising the present invention. In most instances, these promoters may be isolated as convenient restriction digestion fragments suitable for cloning in a selected vector. Alternatively, promoter fragments may be isolated using the polymerase chain

reaction. Cloning of the amplified fragments may be facilitated by incorporating restriction sites at the 5' end of the primers.

The NOI or NOIs may be under the expression control of an expression regulatory element, such as a promoter and enhancer.

Preferably the ischaemic responsive promoter is a tissue restricted ischaemic responsive promoter.

Preferably the tissue restricted ischaemic responsive promoter is a macrophage specific promoter restricted by repression.

Preferably the tissue restricted ischaemic responsive promoter is an endothelium specific promoter.

Preferably the regulated retroviral vector of the present invention is an ILRE regulated retroviral vector.

Preferably the regulated retroviral vector of the present invention is an ILRE regulated lentiviral vector.

Preferably the regulated retroviral vector of the present invention is an autoregulated hypoxia responsive lentiviral vector.

Preferably the regulated retroviral vector of the present invention is regulated by glucose concentration.

For example, the glucose-regulated proteins (grp's) such as grp78 and grp94 are highly conserved proteins known to be induced by glucose deprivation (Attenello and Lee 1984 Science 226 187-190). The grp 78 gene is expressed at low levels in most normal healthy tissues under the influence of basal level promoter elements but has at least two critical "stress inducible regulatory elements" upstream of the TATA element (Attenello 1984 ibid; Gazit *et al* 1995 Cancer Res 55: 1660-1663). Attachment to a truncated 632 base

pair sequence of the 5'end of the grp78 promoter confers high inducibility to glucose deprivation on reporter genes *in vitro* (Gazit *et al* 1995 *ibid*). Furthermore, this promoter sequence in retroviral vectors was capable of driving a high level expression of a reporter gene in tumour cells in murine fibrosarcomas, particularly in central relatively ischaemic/fibrotic sites (Gazit *et al* 1995 *ibid*).

Preferably the regulated retroviral vector of the present invention is a self-inactivating (SIN) vector.

By way of example, self-inactivating retroviral vectors have been constructed by deleting the transcriptional enhancers or the enhancers and promoter in the U3 region of the 3' LTR. After a round of vector reverse transcription and integration, these changes are copied into both the 5' and the 3' LTRs producing a transcriptionally inactive provirus (Yu *et al* 1986 Proc Natl Acad Sci 83: 3194-3198; Dougherty and Temin 1987 Proc Natl Acad Sci 84: 1197-1201; Hawley *et al* 1987 Proc Natl Acad Sci 84: 2406-2410; Yee *et al* 1987 Proc Natl Acad Sci 91: 9564-9568). However, any promoter(s) internal to the LTRs in such vectors will still be transcriptionally active. This strategy has been employed to eliminate effects of the enhancers and promoters in the viral LTRs on transcription from internally placed genes. Such effects include increased transcription (Jolly *et al* 1983 Nucleic Acids Res 11: 1855-1872) or suppression of transcription (Emerman and Temin 1984 Cell 39: 449-467). This strategy can also be used to eliminate downstream transcription from the 3' LTR into genomic DNA (Herman and Coffin 1987 Science 236: 845-848). This is of particular concern in human gene therapy where it is of critical importance to prevent the adventitious activation of an endogenous oncogene.

## RETROVIRAL VECTORS

The regulated retroviral vector of the present invention includes but is not limited to: murine leukemia virus (MLV), human immunodeficiency virus (HIV), equine infectious anaemia virus (EIAV), feline immunodeficiency virus (FIV), caprine encephalitis-arthritis virus (CAEV), mouse mammary tumour virus (MMTV), Rous sarcoma virus (RSV), Fujinami sarcoma virus (FuSV), Moloney murine leukemia virus (Mo-MLV), FBR murine osteosarcoma virus (FBR MSV), Moloney murine sarcoma virus (Mo-MSV),

Abelson murine leukemia virus (A-MLV), Avian myelocytomatosis virus-29 (MC29), and Avian erythroblastosis virus (AEV).

A detailed list of retroviruses may be found in Coffin *et al* ("Retroviruses" 1997 Cold Spring Harbour Laboratory Press Eds: JM Coffin, SM Hughes, HE Varmus pp 758-763).

Preferred vectors for use in accordance with the present invention are retroviral vectors, such as MLV vectors.

Preferably the recombinant retroviral vectors of the present invention are lentiviral vectors, more preferably HIV or EIAV vectors.

LENTIVIRAL VECTORS

The lentiviruses can be divided into primate and non-primate groups. Examples of primate lentiviruses include but are not limited to: the human immunodeficiency virus (HIV), the causative agent of human auto-immunodeficiency syndrome (AIDS), and the simian immunodeficiency virus (SIV). The non-primate lentiviral group includes the prototype "slow virus" visna/maedi virus (VMV), as well as the related caprine arthritis-encephalitis virus (CAEV), equine infectious anaemia virus (EIAV) and the more recently described feline immunodeficiency virus (FIV) and bovine immunodeficiency virus (BIV).

A distinction between the lentivirus family and other types of retroviruses is that lentiviruses have the capability to infect both dividing and non-dividing cells (Lewis *et al* 1992 EMBO. J 11: 3053-3058; Lewis and Emerman 1994 J. Virol. 68: 510-516). In contrast, other retroviruses - such as MLV - are unable to infect non-dividing cells such as those that make up, for example, muscle, brain, lung and liver tissue.

Preferred vectors for use in accordance with the present invention are recombinant retroviral vectors, in particular recombinant lentiviral vectors, in particular minimal lentiviral vectors which are disclosed in WO 99/32646 and in WO98/17815.

VECTOR

As used herein, a "vector" denotes a tool that allows or faciliates the transfer of an entity from one environment to another. In accordance with the present invention, and by way of example, some vectors used in recombinant DNA techniques allow entities, such as a segment of DNA (such as a heterologous DNA segment, such as a heterologous cDNA segment), to be transferred into a target cell. Optionally, once within the target cell, the vector may then serve to maintain the heterologous DNA within the cell or may act as a unit of DNA replication. Examples of vectors used in recombinant DNA techniques include plasmids, chromosomes, artificial chromosomes or viruses.

OPERABLY LINKED

The term "operably linked" denotes a relationship between a regulatory region (typically a promoter element, but may include an enhancer element) and the coding region of a gene, whereby the transcription of the coding region is under the control of the regulatory region.

DERIVABLE

The term "derivable" is used in its normal sense as meaning a nucleotide sequence such as an LTR or a part thereof which need not necessarily be obtained from an vector such as a retroviral vector but instead could be derived therefrom. By way of example, the sequence may be prepared synthetically or by use of recombinant DNA techniques.

VECTOR PARTICLES

In the present invention, several terms are used interchangeably. Thus, "virion", "virus", "viral particle", "retroviral particle", "retrovirus", and "vector particle" mean virus and virus-like particles that are capable of introducing a nucleic acid into a cell through a viral-like entry mechanism. Such vector particles can, under certain circumstances, mediate the transfer of NOIs into the cells they infect. A retrovirus is capable of reverse

transcribing its genetic material into DNA and incorporating this genetic material into a target cell's DNA upon transduction. Such cells are designated herein as "target cells".

A vector particle includes the following components: a retrovirus nucleic acid, which may contain one or more NOIs, a nucleocapsid encapsidating the nucleic acid, the nucleocapsid comprising nucleocapsid protein of a retrovirus, and a membrane surrounding the nucleocapsid.

NUCLEOCAPSID

The term "nucleocapsid" refers to at least the group specific viral core proteins (*gag*) and the viral polymerase (*pol*) of a retrovirus genome. These proteins encapsidate the retrovirus-packagable sequences and themselves are further surrounded by a membrane containing an envelope glycoprotein.

Preferably a high titre retroviral vector is produced using a codon optimised gag and a codon optimised pol or a codon optimised env.

CODON OPTIMISATION

As used herein, the terms "codon optimised" and "codon optimisation" refer to an improvement in codon usage. By way of example, alterations to the coding sequences for viral components may improve the sequences for codon usage in the mammalian cells or other cells which are to act as the producer cells for retroviral vector particle production. This is referred to as "codon optimisation". Many retroviruses, including HIV and other lentiviruses, use a large number of rare codons and by changing these to correspond to commonly used mammalian codons, increased expression of the packaging components in mammalian producer cells can be achieved. Codon usage tables are known in the art for mammalian cells, as well as for a variety of other organisms.

Preferably a high titre lentiviral vector is produced using a codon optimised gag and a codon optimised pol or a codon optimised env.

Preferably a high titre retroviral vector is produced using a modified and/or extended packaging signal.

PACKAGING SIGNAL

As used herein, the term "packaging signal" or "packaging sequence" refers to sequences located within the retroviral genome which are required for insertion of the viral RNA into the viral capsid or particle. Several retroviral vectors use the minimal packaging signal (also referred to as the psi sequence) needed for encapsidation of the viral genome. By way of example, this minimal packaging signal encompasses bases 212 to 563 of the Mo-MLV genome (Mann et al 1983: Cell 33: 153).

As used herein, the term "extended packaging signal" or "extended packaging sequence" refers to the use of sequences around the psi sequence with further extension into the gag gene. The inclusion of these additional packaging sequences may increase the efficiency of insertion of vector RNA into viral particles.

Preferably a high titre lentiviral vector is produced using a modified packaging signal.

Preferably the lentiviral construct is a based on an EIAV vector genome where all the accessory genes are removed except Rev.

ACCESSORY GENES

As used herein, the term "accessory genes" refer to a variety of virally encoded accessory proteins capable of modulating various aspects of retroviral replication and infectivity. These proteins are discussed in Coffin et al (ibid) (Chapters 6 and 7). Examples of accessory proteins in lentiviral vectors include but are not limited to tat, rev, nef, vpr, vpu, vif, vpx. An example of a lentiviral vector useful in the present invention is one which has all of the accessory genes removed except rev.

Preferably the production of lentiviral vector particles is increased by about 10 fold in the presence of EIAV Rev.

ENV

If the retroviral component includes an *env* nucleotide sequence, then all or part of that sequence can be optionally replaced with all or part of another *env* nucleotide sequence such as, by way of example, the amphotropic Env protein designated 4070A or the influenza haemagglutinin (HA) or the vesicular stomatitis virus G (VSV-G) protein. Replacement of the *env* gene with a heterologous *env* gene is an example of a technique or strategy called pseudotyping. Pseudotyping is not a new phenomenon and examples may be found in WO-A-98/05759, WO-A-98/05754, WO-A-97/17457, WO-A-96/09400, WO-A-91/00047 and Mebatsion *et al* 1997 Cell 90, 841-847.

In one preferred aspect, the retroviral vector of the present invention has been pseudotyped. In this regard, pseudotyping can confer one or more advantages. For example, with the lentiviral vectors, the *env* gene product of the HIV based vectors would restrict these vectors to infecting only cells that express a protein called CD4. But if the *env* gene in these vectors has been substituted with *env* sequences from other RNA viruses, then they may have a broader infectious spectrum (Verma and Somia 1997 Nature 389:239-242). By way of example, workers have pseudotyped an HIV based vector with the glycoprotein from VSV (Verma and Somia 1997 *ibid*).

In another alternative, the Env protein may be a modified Env protein such as a mutant or engineered Env protein. Modifications may be made or selected to introduce targeting ability or to reduce toxicity or for another purpose (Valsesia-Wittman *et al* 1996 J Virol 70: 2056-64; Nilson *et al* 1996 Gene Therapy 3: 280-6; Fielding *et al* 1998 Blood 9: 1802 and references cited therein).

TARGET CELL

As used herein the term "target cell" simply refers to a cell which the regulated retroviral vector of the present invention, whether native or targeted, is capable of infecting or transducing.

The lentiviral vector particle according to the invention will be capable of transducing cells which are slowly-dividing, and which non-lentiviruses such as MLV would not be able to efficiently transduce. Slowly-dividing cells divide once in about every three to four days including certain tumour cells. Although tumours contain rapidly dividing cells, some tumour cells especially those in the centre of the tumour, divide infrequently.

Alternatively the target cell may be a growth-arrested cell capable of undergoing cell division such as a cell in a central portion of a tumour mass or a stem cell such as a haematopoietic stem cell or a CD34-positive cell.

As a further alternative, the target cell may be a precursor of a differentiated cell such as a monocyte precursor, a CD33-positive cell, or a myeloid precursor.

As a further alternative, the target cell may be a differentiated cell such as a neuron, astrocyte, glial cell, microglial cell, macrophage, monocyte, epithelial cell, endothelial cell, hepatocyte, spermatocyte, spermatid or spermatozoa.
Target cells may be transduced either *in vitro* after isolation from a human individual or may be transduced directly *in vivo*.

NOI

In accordance with the present invention, it is possible to manipulate the viral genome or the regulated retroviral vector nucleotide sequence, so that viral genes are replaced or supplemented with one or more NOIs which may be heterologous NOIs.

The term "heterologous" refers to a nucleic acid sequence or protein sequence linked to a nucleic acid or protein sequence which it is not naturally linked.

With the present invention, the term NOI (i.e. nucleotide sequence of interest) includes any suitable nucleotide sequence, which need not necessarily be a complete naturally occurring DNA sequence. Thus, the DNA sequence can be, for example, a synthetic DNA sequence, a recombinant DNA sequence (i.e. prepared by use of recombinant DNA techniques), a cDNA sequence or a partial genomic DNA sequence, including

combinations thereof. The DNA sequence need not be a coding region. If it is a coding region, it need not be an entire coding region. In addition, the DNA sequence can be in a sense orientation or in an anti-sense orientation. Preferably, it is in a sense orientation. Preferably, the DNA is or comprises cDNA.

The NOI(s) may be any one or more of selection gene(s), marker gene(s) and therapeutic gene(s). As used herein, the term "selection gene" refers to the use of a NOI which encodes a selectable marker which may have an enzymatic activity that confers resistance to an antibiotic or drug upon the cell in which the selectable marker is expressed.

SELECTABLE MARKERS

Many different selectable markers have been used successfully in retroviral vectors. These are reviewed in "Retroviruses" (1997 Cold Spring Harbour Laboratory Press Eds: JM Coffin, SM Hughes, HE Varmus pp 444) and include, but are not limited to, the bacterial neomycin (neo) and hygromycin phosphotransferase genes which confer resistance to G418 and hygromycin respectively; a mutant mouse dihydrofolate reductase gene which confers resistance to methotrexate; the bacterial *gpt* gene which allows cells to grow in medium containing mycophenolic acid, xanthine and aminopterin; the bacterial *hisD* gene which allows cells to grow in medium without histidine but containing histidinol; the multidrug resistance gene (*mdr*) which confers resistance to a variety of drugs; and the bacterial genes which confer resistance to puromycin or phleomycin. All of these markers are dominant selectable and allow chemical selection of most cells expressing these genes. Other selectable markers are not dominant in that their use must be in conjunction with a cell line that lacks the relevant enzyme activity. Examples of non-dominant selectable markers include the thymidine kinase (*tk*) gene which is used in conjunction with *tk* cell lines.

Particularly preferred markers are blasticidin and neomycin, optionally operably linked to a thymidine kinase coding sequence typically under the transcriptional control of a strong viral promoter such the SV40 promoter.

NOIs WITH THERAPEUTIC AND/OR DIAGNOSTIC APPLICATIONS

In accordance with the present invention, suitable NOI sequences include those that are of therapeutic and/or diagnostic application such as, but are not limited to: sequences encoding cytokines, chemokines, hormones, antibodies, engineered immunoglobulin-like molecules, a single chain antibody, fusion proteins, enzymes, immune co-stimulatory molecules, immunomodulatory molecules, anti-sense RNA, a transdominant negative mutant of a target protein, a toxin, a conditional toxin, an antigen, a tumour suppressor protein and growth factors, membrane proteins, vasoactive proteins and peptides, anti-viral proteins and ribozymes, and derivatives therof (such as with an associated reporter group). When included, such coding sequences may be typically operatively linked to a suitable promoter, which may be a promoter driving expression of a ribozyme(s), or a different promoter or promoters, such as in one or more specific cell types.

NOIs FOR TREATING CANCER

Suitable NOIs for use in the invention in the treatment or prophylaxis of cancer include NOIs encoding proteins which: destroy the target cell (for example a ribosomal toxin), act as: tumour suppressors (such as wild-type p53); activators of anti-tumour immune mechanisms (such as cytokines, co-stimulatory molecules and immunoglobulins); inhibitors of angiogenesis; or which provide enhanced drug sensitivity (such as pro-drug activation enzymes); indirectly stimulate destruction of target cell by natural effector cells (for example, strong antigen to stimulate the immune system or convert a precursor substance to a toxic substance which destroys the target cell (for example a prodrug activating enzyme).

PRO-DRUG ACTIVATING ENZYMES

Examples of prodrugs include but are not limited to etoposide phosphate (used with alkaline phosphatase; 5-fluorocytosine (with cytosine deaminase); Doxorubin-N-p-hydroxyphenoxyacetamide (with Penicillin-V-Amidase); Para-N-bis (2-chloroethyl)aminobenzoyl glutamate (with Carboxypeptidase G2); Cephalosporin nitrogen mustard carbamates (with B-lactamase); SR4233 (with p450 reductase);

Ganciclovir (with HSV thymidine kinase); mustard pro-drugs with nitroreductase and cyclophosphamide or ifosfamide (with cytochrome p450).

NOIs FOR TREATING HEART DISEASE

Suitable NOIs for use in the treatment or prevention of ischaemic heart disease include NOIs encoding plasminogen activators. Suitable NOIs for the treatment or prevention of rheumatoid arthritis or cerebral malaria include genes encoding anti-inflammatory proteins, antibodies directed against tumour necrosis factor (TNF) alpha, and anti-adhesion molecules (such as antibody molecules or receptors specific for adhesion molecules).

BYSTANDER EFFECT

The expression products encoded by the NOIs may be proteins which are secreted from the cell. Alternatively the NOI expression products are not secreted and are active within the cell. In either event, it is preferred for the NOI expression product to demonstrate a bystander effector or a distant bystander effect; that is the production of the expression product in one cell leading to the killing of additional, related cells, either neighbouring or distant (e.g. metastatic), which possess a common phenotype. Encoded proteins could also destroy bystander tumour cells (for example with secreted antitumour antibody-ribosomal toxin fusion protein), indirectly stimulated destruction of bystander tumour cells (for example cytokines to stimulate the immune system or procoagulant proteins causing local vascular occlusion) or convert a precursor substance to a toxic substance which destroys bystander tumour cells (eg an enzyme which activates a prodrug to a diffusible drug). Also, the delivery of NOI(s) encoding antisense transcripts or ribozymes which interfere with expression of cellular genes for tumour persistence (for example against aberrant *myc* transcripts in Burkitts lymphoma or against *bcr-abl* transcripts in chronic myeloid leukemia. The use of combinations of such NOIs is also envisaged.

CYTOKINES

The NOI or NOIs of the present invention may also comprise one or more cytokine-encoding NOIs. Suitable cytokines and growth factors include but are not limited to: ApoE, Apo-SAA, BDNF, Cardiotrophin-1, EGF, ENA-78, Eotaxin, Eotaxin-2, Exodus-2, FGF-acidic, FGF-basic, fibroblast growth factor-10 (Marshall 1998 Nature Biotechnology 16: 129).FLT3 ligand (Kimura *et al* (1997), Fractalkine (CX3C), GDNF, G-CSF, GM-CSF, GF-β1, insulin, IFN-γ, IGF-I, IGF-II, IL-1α, IL-1β, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8 (72 a.a.), IL-8 (77 a.a.), IL-9, IL-10, IL-11, IL-12, IL-13, IL-15, IL-16, IL-17, IL-18 (IGIF), Inhibin α, Inhibin β, IP-10, keratinocyte growth factor-2 (KGF-2), KGF, Leptin, LIF, Lymphotactin, Mullerian inhibitory substance, monocyte colony inhibitory factor, monocyte attractant protein (Marshall 1998 *ibid*), M-CSF, MDC (67 a.a.), MDC (69 a.a.), MCP-1 (MCAF), MCP-2, MCP-3, MCP-4, MDC (67 a.a.), MDC (69 a.a.), MIG, MIP-1α, MIP-1β, MIP-3α, MIP-3β, MIP-4, myeloid progenitor inhibitor factor-1 (MPIF-1), NAP-2, Neurturin, Nerve growth factor, β-NGF, NT-3, NT-4, Oncostatin M, PDGF-AA, PDGF-AB, PDGF-BB, PF-4, RANTES, SDF1α, SDF1β, SCF, SCGF, stem cell factor (SCF), TARC, TGF-α, TGF-β, TGF-β2, TGF-β3, tumour necrosis factor (TNF), TNF-α, TNF-β, TNIL-1, TPO, VEGF, GCP-2, GRO/MGSA, GRO-β, GRO-γ, HCC1, 1-309.

The NOI or NOIs may be under the expression control of an expression regulatory element, such as a promoter and/or a promoter enhancer as known as "responsive elements" in the present invention.

VIRAL DELIVERY SYSTEMS

When the regulated retroviral vector particles are used to transfer NOIs into cells which they transduce, such vector particles also designated "viral delivery systems" or "retroviral delivery systems". Viral vectors, including retroviral vectors, have been used to transfer NOIs efficiently by exploiting the viral transduction process. NOIs cloned into the retroviral genome can be delivered efficiently to cells susceptible to transduction by a retrovirus. Through other genetic manipulations, the replicative capacity of the retroviral

genome can be destroyed. The vectors introduce new genetic material into a cell but are unable to replicate.

The regulated retroviral vector of the present invention can be delivered by viral or non-viral techniques. Non-viral delivery systems include but are not limted to DNA transfection methods. Here, transfection includes a process using a non-viral vector to deliver a gene to a target mammalian cell.

Typical transfection methods include electroporation, DNA biolistics, lipid-mediated transfection, compacted DNA-mediated transfection, liposomes, immunoliposomes, lipofectin, cationic agent-mediated, cationic facial amphiphiles (CFAs) (Nature Biotechnology 1996 14; 556), multivalent cations such as spermine, cationic lipids or polylysine, 1, 2,-bis (oleoyloxy)-3-(trimethylammonio) propane (DOTAP)-cholesterol complexes (Wolff and Trubetskoy 1998 Nature Biotechnology 16: 421) and combinations thereof.

Viral delivery systems include but are not limited to adenovirus vector, an adeno-associated viral (AAV) vector, a herpes viral vector, a retroviral vector, a lentiviral vector, or a baculoviral vector. These viral delivery systems may be configured as a split-intron vector. A split intron vector is described in WO 99/15683.

Other examples of vectors include *ex vivo* delivery systems, which include but are not limited to DNA transfection methods such as electroporation, DNA biolistics, lipid-mediated transfection, compacted DNA-mediated transfection.

The vector may be a plasmid DNA vector. Alternatively, the vector may be a recombinant viral vector. Suitable recombinant viral vectors include adenovirus vectors, adeno-associated viral (AAV) vectors, Herpes-virus vectors, or retroviral vectors, lentiviral vectors or a combination of adenoviral and lentiviral vectors. In the case of viral vectors, gene delivery is mediated by viral infection of a target cell.

If the features of adenoviruses are combined with the genetic stability of retro/lentiviruses then essentially the adenovirus can be used to transduce target cells to become transient retroviral producer cells that could stably infect neighbouring cells.

PHARMACEUTICAL COMPOSITION

The present invention also provides a pharmaceutical composition for treating an individual by gene therapy, wherein the composition comprises a therapeutically effective amount of a regulated retroviral vector according to the present invention. The pharmaceutical composition may be for human or animal usage. Typically, a physician will determine the actual dosage which will be most suitable for an individual subject and it will vary with the age, weight and response of the particular patient.

The composition may optionally comprise a pharmaceutically acceptable carrier, diluent, excipient or adjuvant. The choice of pharmaceutical carrier, excipient or diluent can be selected with regard to the intended route of administration and standard pharmaceutical practice. The pharmaceutical compositions may comprise as - or in addition to - the carrier, excipient or diluent any suitable binder(s), lubricant(s), suspending agent(s), coating agent(s), solubilising agent(s), and other carrier agents that may aid or increase the viral entry into the target site (such as for example a lipid delivery system).

Where appropriate, the pharmaceutical compositions can be administered by any one or more of: minipumps, inhalation, in the form of a suppository or pessary, topically in the form of a lotion, solution, cream, ointment or dusting powder, by use of a skin patch, orally in the form of tablets containing excipients such as starch or lactose, or in capsules or ovules either alone or in admixture with excipients, or in the form of elixirs, solutions or suspensions containing flavouring or colouring agents, or they can be injected parenterally, for example intracavernosally, intravenously, intramuscularly or subcutaneously. For parenteral administration, the compositions may be best used in the form of a sterile aqueous solution which may contain other substances, for example enough salts or monosaccharides to make the solution isotonic with blood. For buccal or sublingual administration the compositions may be administered in the form of tablets or lozenges which can be formulated in a conventional manner.

DISORDERS

The present invention is believed to have a wide therapeutic applicability - depending on *inter alia* the selection of the one or more NOIs.

For example, the present invention may be useful in the treatment of the disorders listed in WO-A-98/05635. For ease of reference, part of that list is now provided: cancer, inflammation or inflammatory disease, dermatological disorders, fever, cardiovascular effects, haemorrhage, coagulation and acute phase response, cachexia, anorexia, acute infection, HIV infection, shock states, graft-versus-host reactions, autoimmune disease, reperfusion injury, meningitis, migraine and aspirin-dependent anti-thrombosis; tumour growth, invasion and spread, angiogenesis, metastases, malignant, ascites and malignant pleural effusion; cerebral ischaemia, ischaemic heart disease, osteoarthritis, rheumatoid arthritis, osteoporosis, asthma, multiple sclerosis, neurodegeneration, Alzheimer's disease, atherosclerosis, stroke, vasculitis, Crohn's disease and ulcerative colitis; periodontitis, gingivitis; psoriasis, atopic dermatitis, chronic ulcers, epidermolysis bullosa; corneal ulceration, retinopathy and surgical wound healing; rhinitis, allergic conjunctivitis, eczema, anaphylaxis; restenosis, congestive heart failure, endometriosis, atherosclerosis or endosclerosis.

In addition, or in the alternative, the present invention may be useful in the treatment of disorders listed in WO-A-98/07859. For ease of reference, part of that list is now provided: cytokine and cell proliferation/differentiation activity; immunosuppressant or immunostimulant activity (e.g. for treating immune deficiency, including infection with human immune deficiency virus; regulation of lymphocyte growth; treating cancer and many autoimmune diseases, and to prevent transplant rejection or induce tumour immunity); regulation of haematopoiesis, e.g. treatment of myeloid or lymphoid diseases; promoting growth of bone, cartilage, tendon, ligament and nerve tissue, e.g. for healing wounds, treatment of burns, ulcers and periodontal disease and neurodegeneration; inhibition or activation of follicle-stimulating hormone (modulation of fertility); chemotactic/chemokinetic activity (e.g. for mobilising specific cell types to sites of injury or infection); haemostatic and thrombolytic activity (e.g. for treating haemophilia and stroke); antiinflammatory activity (for treating e.g. septic shock or Crohn's disease); as

antimicrobials; modulators of e.g. metabolism or behaviour; as analgesics; treating specific deficiency disorders; in treatment of e.g. psoriasis, in human or veterinary medicine.

In addition, or in the alternative, the present invention may be useful in the treatment of disorders listed in WO-A-98/09985. For ease of reference, part of that list is now provided: macrophage inhibitory and/or T cell inhibitory activity and thus, anti-inflammatory activity; anti-immune activity, i.e. inhibitory effects against a cellular and/or humoral immune response, including a response not associated with inflammation; inhibit the ability of macrophages and T cells to adhere to extracellular matrix components and fibronectin, as well as up-regulated fas receptor expression in T cells; inhibit unwanted immune reaction and inflammation including arthritis, including rheumatoid arthritis, inflammation associated with hypersensitivity, allergic reactions, asthma, systemic lupus erythematosus, collagen diseases and other autoimmune diseases, inflammation associated with atherosclerosis, arteriosclerosis, atherosclerotic heart disease, reperfusion injury, cardiac arrest, myocardial infarction, vascular inflammatory disorders, respiratory distress syndrome or other cardiopulmonary diseases, inflammation associated with peptic ulcer, ulcerative colitis and other diseases of the gastrointestinal tract, hepatic fibrosis, liver cirrhosis or other hepatic diseases, thyroiditis or other glandular diseases, glomerulonephritis or other renal and urologic diseases, otitis or other oto-rhino-laryngological diseases, dermatitis or other dermal diseases, periodontal diseases or other dental diseases, orchitis or epididimo-orchitis, infertility, orchidal trauma or other immune-related testicular diseases, placental dysfunction, placental insufficiency, habitual abortion, eclampsia, pre-eclampsia and other immune and/or inflammatory-related gynaecological diseases, posterior uveitis, intermediate uveitis, anterior uveitis, conjunctivitis, chorioretinitis, uveoretinitis, optic neuritis, intraocular inflammation, e.g. retinitis or cystoid macular oedema, sympathetic ophthalmia, scleritis, retinitis pigmentosa, immune and inflammatory components of degenerative fondus disease, inflammatory components of ocular trauma, ocular inflammation caused by infection, proliferative vitreo-retinopathies, acute ischaemic optic neuropathy, excessive scarring, e.g. following glaucoma filtration operation, immune and/or inflammation reaction against ocular implants and other immune and inflammatory-related ophthalmic diseases, inflammation associated with autoimmune diseases or conditions or disorders

where, both in the central nervous system (CNS) or in any other organ, immune and/or inflammation suppression would be beneficial, Parkinson's disease, complication and/or side effects from treatment of Parkinson's disease, AIDS-related dementia complex HIV-related encephalopathy, Devic's disease, Sydenham chorea, Alzheimer's disease and other degenerative diseases, conditions or disorders of the CNS, inflammatory components of stokes, post-polio syndrome, immune and inflammatory components of psychiatric disorders, myelitis, encephalitis, subacute sclerosing pan-encephalitis, encephalomyelitis, acute neuropathy, subacute neuropathy, chronic neuropathy, Guillaim-Barre syndrome, Sydenham chora, myasthenia gravis, pseudo-tumour cerebri, Down's Syndrome, Huntington's disease, amyotrophic lateral sclerosis, inflammatory components of CNS compression or CNS trauma or infections of the CNS, inflammatory components of muscular atrophies and dystrophies, and immune and inflammatory related diseases, conditions or disorders of the central and peripheral nervous systems, post-traumatic inflammation, septic shock, infectious diseases, inflammatory complications or side effects of surgery, bone marrow transplantation or other transplantation complications and/or side effects, inflammatory and/or immune complications and side effects of gene therapy, e.g. due to infection with a viral carrier, or inflammation associated with AIDS, to suppress or inhibit a humoral and/or cellular immune response, to treat or ameliorate monocyte or leukocyte proliferative diseases, e.g. leukaemia, by reducing the amount of monocytes or lymphocytes, for the prevention and/or treatment of graft rejection in cases of transplantation of natural or artificial cells, tissue and organs such as cornea, bone marrow, organs, lenses, pacemakers, natural or artificial skin tissue.

## INTRODUCTION TO THE EXAMPLES SECTION AND THE FIGURES

The present invention will now be described only by way of example in which reference is made to the following Figures:

Figure 1 shows an MLV-based transduction method using a Cre/LoxP system as described by Vanin *et al ibid* (1997);

Figure 2 shows an EIAV-based transduction method using a Cre/Lox system;

Figure 3 shows an MLV SIN vector construct transduction method with an EIAV/HIV genome insertion using a Cre/Lox system;

Figure 4 shows an MLV-based transduction method with HRE 3'LTR using a Cre/Lox P system;

Figure 5 shows an MLV-based transduction method for MLV SIN vector production using a Cre/Lox P system;

Figure 6 shows an MLV-based transduction method with integration of a complete second genome construct using a Cre/LoxP system;

Figure 7 shows the basis molecular organisation of an RNA genome and a proviral DNA genome;

Figure 8 shows a schematic diagram of pTrap2 and pONY8z-loxP plasmids;

Figure 9 shows an overall summary of the recombinase method;

Figure 10a shows a FACS analysis of EV1 packaging cells prior to transduction with Trap2 vector;

Figure 10b shows FACS analysis of EV1 packaging cell line transduced with Trap2 at an MOI of 0.3. A 5% top slice of the highest expressers was carried out;

Figure 11 shows a validation of the method for quantitation of GFP mRNA, relative to β-actin. A titration of the total RNA from EV1 clone A was used. The difference in Ct values between the two assays is shown on the y axis. The magnitude of the gradient must be <0.1 for the method to be valid. The gradient is 0.077. so the method is suitable;

Figure 12 shows the quantitation of GFP mRNA relative to control β-actin mRNA. EV2 TD cells are transduced with Trap2 at an MOI of 0.3 and are the calibrator sample with the ratio designated 1.0;

Figure 13 shows FACS analysis of EV1 clone A:

Figure 13A shows original GFP expression of the clone;

Figure 13B shows GFP expression 7 days after transfection with Cre recombinase (pBS185). Excision frequency is 64%:

Figure 13C shows recombined clone 4 identified as being negative for GFP:

Figure 14 shows lacZ expression of transfected cells with and without the addition of the Cre recombinase (pBS185). Figure 14 shows EV1A4 and EV2D4 clones with and without the addition of Cre recombinase (pBS185). The efficiency of the insertion event was estimated to be about 12% by computer image analysis;

Figure 15 shows the structure of pONY8.1Z MLVHyb;

Figure 16 shows the alignment of leader and gag regions present in vectors pONY4Z, 8Z and ATG mutated 8Z vector. The latter is referred to as pONY8ZA. The sequences aligned are from the NarI site in the leader to the XbaI site between the EIAV gag sequence and the CMV promoter. Sequences in the leader are shown in italic and a space is present upstream of the position of the gag ATG; and

Figure 17 shows a schematic representation of the structure of pONY 8.3G +/- vector genome plasmids.

## EXAMPLES

EXAMPLE 1

Vanin *et al* (*ibid*) describe a recombinase system whereby an initial retroviral transduction event introduces retroviral LTRs and expressed gene/s flanked by two recombinase target sites (exemplified by loxP) into a cell line. Stable transduced cell lines are selected by resistance to the antibiotic neomycin and screened for high expression of

the expressed gene(s) (see Figure 1). Such cell lines (Cell Line 1) contain retroviral insertions in integration sites that support high level expression from the retroviral genome.

The next step involves the transfection of the relevant recombinase expression construct (exemplified here by Cre recombinase) into the identified high expressing cell line. The expressed gene(s) is/are excised and a single loxP site is retained in the construct (Cell Line 2). In this instance, thymidine kinase gene (tk) is used as a negative selectible marker in combination with the drug, gancyclovir. The final step involves the re-insertion of a therapeutic or marker gene of choice into the single loxP site via a Cre-assisted mechanism. Cell lines are identified that have been successfully recombined (Cell Line 3) and they will produce retroviruses at the same titre as the parental Cell Line 1.

## EXAMPLE 2

Figure 2 and Figure 3 describe the production of EIAV or HIV high titre transduced producer cell lines.

Figure 2 shows a minimal EIAV genome construct with the 3' U3 sequences replaced by a strong constitutive promoter, CMV. A reporter gene such as blasticidin resistance gene (*bsr*) is flanked by loxP sites. Virus is made in a transient system and is transduced into an EIAV producer cell line and clones identified that maximally express the blast marker gene. A line is chosen (termed Cell Line 1) and the marker gene is excised by a Cre recombinase-assisted excision event, generating Cell Line 2.

Construct B comprises two loxP sites which flank an internal expression cassette and also the native EIAV 3' LTR. Therefore, this construct is recombined into the cell line such that the 5' R and U5 sequences are inherited from the packaging cell line, whereas the 3' LTR sequences are wholly derived from the recombined construct. The 3' LTR from Cell Line 2 is present downstream of the functional EIAV genome expression construct. This CMV-R-U5 module is still transcriptionally active but expression is directed away from the EIAV genome.

Figure 3 shows a further aspect of the invention. Construct C is based on an MLV SIN vector, with a deletion in the 3' U3 sequences. The cassette includes an internal CMV promoter linked to EIAV R and U5 sequences. This is followed by a blasticidin resistance gene (*bsr*) flanked by two loxP sites. Virus is made in a transient transfection system and the genome is transduced into a packaging line. Blast-resistant clones are identified and the highest expressing line is chosen for further analysis. This line is transfected with Cre recombinase and the blast gene is excised. The last step involves the insertion of construct B into the single loxP site. Once again, a complete EIAV 3' LTR is introduced into the producer cell line. This leads to a CMV-driven EIAV genome expression cassette with the EIAV 3' LTR still located at the 3' end of the genome. Transcriptionally quiescent MLV SIN LTRs flanks these EIAV sequences.

**EXAMPLE 3**

Figure 4 shows an additional aspect of the invention. Construct D is an MLV-based vector with a CMV promoter in the 3' LTR in place of the U3 sequences. Virus is made in a transient system and is transduced into a packaging cell line as described previously. The neo and TK genes are excised by the action of Cre recombinase and construct E is recombined into the single loxP target sequence. The modified MLV 3' LTR including the HRE or similarly regulated system is transferred into the packaging cell line by the recombinase mechanism. Therefore, the 5' R and U5 sequences are inherited from the producer cell line whereas the therapeutic and marker gene/s and regulated 3' LTR is inherited from construct E. The final producer cell line is constitutively driven by the 5' CMV promoter and will produce high titre retroviral vectors which are regulated in the transduced target cells. This approach avoids the derivation of low titre transfected producer cell lines or the use of hypoxic conditions or chemical mimics for production from traditionally derived transduced producer lines.

**EXAMPLE 4**

Figure 5 shows yet another aspect of the invention. Construct D is an MLV-based vector with a CMV promoter in the 3' LTR as previously described. The same process is carried

out as shown in Figure 4 until the final recombination is performed. Construct F contains a deletion in U3 sequences in the 3' LTR and an internal expression cassette comprising a promoter and gene sequences. The final cell line containing the Cre-mediated recombination will be CMV-driven and will constitutively produce high titre MLV SIN vectors. Previously, SIN vectors have not been amenable to production by stable cell line producer technology. Instead they have been prepared using transfection-based transient expression systems.

## EXAMPLE 5

Figure 6 shows an MLV-based transduction method with integration of complete second genome construct by Cre/LoxP system. In this approach, construct 1 is called TRAP1) is an MLV vector construct containing an internal CMV promoter operably linked to a marker gene (a truncated form of the human low affinity nerve growth factor receptor, called LNGFR). The enhancer elements in the 3' U3 sequence have been excised and replaced by a 34bp loxP site. Virus stocks are prepared in a transient system and the TRAP1 genome is stably transduced into packaging cell lines.

The modified 3'U3 sequences, including the lox P sequence, is copied from the 3'LTR position to the 5'LTR, such that there is little 5' promoter activity. Cell lines are screened for high levels of expression of LNGFR protein by fluorescent activated cell sorter (FACS) analysis and clonal lines are derived by standard techniques. A Cre recombinase expression plasmid is transfected into the derived cell line to excise all sequences between the two loxP sites. Next, cells are negatively selected by FACS for absence of LNGFR expression and clonal lines are derived by standard techniques. Construct 2 in this example comprises a complete HIV or EIAV or also MLV retroviral genome, which is flanked by two minimal 34bp loxP recombinase sites. A strong constitutive promoter such as CMV directs transcription of the genome. On transfection of plasmid 2 and Cre expression plasmid, the complete lentivirus vector or MLV vector genome is inserted in the producer cell line. These sequences are flanked to the 5' by a small portion of MLV U3 sequence and a loxP site and to the 3' by the second loxP site, enhancerless-U3 sequences, R and U5 derived from the MLV construct 1.

**Derivation of Plasmid TRAP1 (Figure 6 - Construct 1)**

Oligonucleotides VSAT129 and VSAT130 were synthesised which correspond to the minimal 34bp loxP sites and contain a 5' overhang for NheI and a 3' overhang for XbaI. The sequences 5' to 3' are as follows: VSAT129 (CTAGCATAACTTCGTATA ATGTATGCTATACGAAGTTATT) (SEQ ID No 49) and VSAT130 (CTAGAATAACTTCGTATAGC ATACATTATACGAAGTTATG) (SEQ ID No 50). The two oligonucleotides were treated with T4 polynucleotide kinase and were heated to 95°C for 5 minutes, before gradual cooling to room temperature. The annealed and kinased oligos were ligated to a 2,830 bp NheI/XbaI fragment from LTR plasmid (SEQ ID No 59). Fragments were ligated and correct clones of LTRloxP were identified by sequence analysis. Plasmid LTRloxP was then digested with NheI and ScaI and a 2,185bp fragment was prepared for following cloning steps.

Plasmids TRAP1 and TRAP1G were derived from LTRloxP and the MLV genome CGCLNGFR (encodes GFP and LNGFR from an internal CMV promoter – see SEQ ID No 57). However, the GFP gene was excised by EcoRI/BsmI digestion and the 6,796bp fragment was filled in by T4 DNA polymerase and re-ligated, in order to generate plasmid CXCLNGFR. Plasmid TRAP1 was generated by ligation of a 2,185bp NheI/ScaI fragment from LTRloxP (see SEQ ID No 58) to a 4,426bp NheI/ScaI fragment from CXCLNGFR. Plasmid TRAP1G was generated by ligation of a 2,185bp NheI/ScaI fragment from LTRloxP to a 5,179bp NheI/ScaI fragment from CGCLNGFR.

**Derivation of Plasmid pONY8z-lox (Figure 6 - Construct 2)**

In this example, the retroviral genome inserted into the loxP site in Figure 6 was based on the EIAV vector genome, pONY8z (for preparation see pONY8.0Z construction below). pONY8z was cut with SnaBI and NruI, and the 4358bp fragment purified and self-ligated to form pONY8z-shuttle. This plasmid has unique 5' sites (DraIII and BglII) and unique 3' sites (PvuII and BspLUII). Oligonucleotides encoding the 34bp loxP sites were inserted with suitable base pair overhangs at the unique 5' DraIII site and then the unique 3' BspLUII, to generate plasmid pONY-8z-shuttleloxP.

Plasmid pONY8z-loxP was made as follows. Plasmid pONY-8z-shuttleloxP was digested with BsrG I and NspV, and the 3670bp fragment was purified as the vector fragment. The insert for ligation to this fragment was derived from pONY8z by partial digestion with BsrGI (two sites) followed by digestion with NspV. A 7,328bp fragment was purified and ligated to the 3670bp fragment described above.

The Cre recombinase plasmid as used in this system is pBS185 (Gibco).

**EXAMPLE 6**

We constructed an MLV self-inactivating (SIN) vector called pTrap2 (see SEQ ID No 56) by replacing the 3' U3 NheI-XbaI fragment with a 34-bp loxP sequence. The vector transcribes the marker gene GFP from an internal CMV promoter. Trap2 vector was used to transduce EIAV packaging cell lines EV1 and EV2. The EV cell lines are based on human TE671 cells and express EIAV gag/pol proteins and VSV-G envelope, regulated by a temperature-sensitive switch. High expressing clones of transduced EV1 and EV2 cells were identified by FACS analysis for GFP. Individual clones expressing high levels of GFP were then selected. The GFP expression cassette was excised following transient transfection with a Cre recombinase expression plasmid. The derived cell line, EV-loxP, contains a single loxP site and minimal sequences derived from the MLV construct pTrap2. An EIAV genome was engineered to contain loxP sites flanking the entire vector genome.

This genome construct and Cre recombinase were co-transfected into EV-loxP. Stable cell lines expressing lacZ were selected by FACS and cell lines were cloned by limiting dilution. Therefore, we have introduced an entire EIAV genome expression cassette into a single loxP site. This site was previously identified by MLV transduction as highly permissive for transgene expression. A 5' CMV promoter transcribes the lentiviral genome in the producer cell line but the expression site was originally identified by MLV transduction. This method is adaptable to the generation of transduced producer cell lines for other lentiviral vector systems.

## Materials and Methods

**Vector construction:** Plasmid pTrap2 was made as follows: A plasmid containing a single MLV LTR plasmid (LTRplasmid – SEQ ID No 59) was digested with *Nhe*I and *Xba*I and a 34 bp minimal loxP site was introduced with relevant sticky ends. This insertion step removes the MLV U3 enhancer elements which lie within the excised *Nhe*I-*Xba*I fragment. The LTR-loxP plasmid was linearised by digestion with *Nhe*I and was ligated to a 6.8kb *Nhe*I fragment from the MLV construct CZCG (See SEQ ID No 55). This construct expresses lacZ from the 5' U3 promoter and GFP from an internal CMV promoter. The resulting pTrap2 construct is shown in Figure 8.

The EIAV genome construct pONY8.0Z and pONY8.1Z were prepared as follows:

### pONY8.0Z construction

pONY8.0Z was derived from pONY4.0Z (see WO 99/32646) by introducing mutations which 1) prevented expression of TAT by an 83nt deletion in the exon 2 of tat ) prevented S2 ORF expression by a 51nt deletion 3) prevented REV expression by deletion of a single base within exon 1 of rev and 4) prevented expression of the N-terminal portion of gag by insertion of T in ATG start codons, thereby changing the sequence to ATTG from ATG. With respect to the wild type EIAV sequence Acc. No. U01866 these correspond to deletion of nt 5234-5316 inclusive, nt 5346-5396 inclusive and nt 5538. The insertion of T residues was after nt 526 and 543.

### pONY8.1Z construction

pONY8.1Z was obtained directly from pONY8.0Z by digestion with SalI and partial digestion with SapI. Following restriction the overhanging termini of the DNA were made blunt ended by treatment with T4 DNA polymerase. The resulting DNA was then religated. This manipulation results in a deletion of sequence between the LacZ reporter gene and just upstream of the 3'PPT. The 3' border of the deletion is nt 7895 with respect to wild type EIAV, Acc. No. U01866. Thus pONY8.1Z does not contain sequences corresponding to the EIAV RREs.

Plasmid pONY8z was linearised by *Bgl*II, and a single loxP site was cloned into *Bgl*II, immediately upstream of the 5' CMV promoter, to produce pONY8z-loxP. Plasmids pONY3.2iresHYG and pHCMV-VSVG were used in the derivation of cell lines EV1 and EV2. The plasmid pONY3.2iresHYG was constructed as follows:

**pONY3.2IREShyg**

pONY3.IREShyg was derived from pONY3.2. pONY3.2 is a derivative of pONY3.1 in which expression of TAT and S2 are ablated by an 83nt deletion in the exon 2 of tat a 51nt deletion in S2 ORF. With respect to the wild type EIAV sequence Acc. No. U01866 these correspond to deletion of nt 5234-5316 inclusive and nt 5346-5396 inclusive. This fragment was introduced into the expression vector pHORSE IRES hyg which was made as follows. pHORSE (see WO 99/32646) was cut with SnaBI and NotI which excises a fragment running from the CMV promoter through EIAV gag/pol and introduced into pIRES1hyg (Clontech) digested with the same enzymes. This plasmid was then cut with Sse8387I and BstEII and then ligated with the Sse8387I to BstEII fragment from pONY3.2. The sequence of the plasmid is set out in SEQ ID No 51.

**Virus Production**

Transient MLV vector preparations pseudotyped with RD114 cat endogenous envelope were made as described previously (Soneoka et al., 1995). EIAV vector was harvested from confluent monolayers following 3 days induction of VSV-G expression at 32°C. MLV vector preparations were titred in triplicate on HT1080 fibrosarcoma cells. EIAV vector preparations were titred by GFP and lacZ on D17 dog osteosarcoma cells.

**Flow cytometry of b-galactosidase and GFP activity:**

$1.5 \times 10^5$ cells from a 12-well plate were analysed for lacZ expression using the FluoReporter lacZ Flow Cytometry kit (Molecular Probes). GFP expression was also directly assessed using the FACSCalibur flow cytometer (Beckton Dickinson).

**Transfection methods**

Calcium phosphate transfections were carried out using the Profection kit (Promega) according to manufacturer's instructions.

**Results**

Figure 8 shows a schematic diagram of pTrap2 and pONY8z-loxP, plasmids used in this study.

**Introduction of Trap2 genome into EV1 and EV2**

An overall summary of the process described here is given in Figure 9. Trap2 MLV vector was made in a transient system with the amphotropic 4070A envelope. It gave a GFP titre of $1.7x10^6$ T.U. per ml. Trap2 vector however also gave a lacZ titre of $9.4x10^5$ T.U. per ml. This shows that replacement of the *Nhe*I-*Xba*I fragment from the MLV U3 region with loxP does not completely inactivate the MLV U3 promoter. Therefore Trap2, as constructed, is a partial SIN vector.

EV1 and EV2 cells were transduced with Trap2 vector at a multiplicity of infection (MOI) of 0.3. This was done to insert single copies of the MLV genome into the packaging lines.

**Derivation of high expressers of GFP marker gene**

Transduced EV1 and EV2 cells were analysed by FACS (see Figure 10) and the top 5% of GFP expressing cells were sorted and expanded. Clonal lines were derived by limiting dilution and four clones of EV1 and EV2 were chosen by visual inspection.

A quantitative TaqMan RT-PCR reaction was established in order to identify which of the four clones of EV1 and EV2 were the highest expressors of GFP mRNA. Total RNA was analysed by RT-PCR for GFP and β-actin. Quantitation was calculated by direct comparison of the Ct values (Cycle threshold). This was possible as it was proved that the two individual RT-PCR reactions are of similar efficiency (see Figure 11). By identifying an

optimal chromosomal location for GFP transgene expression. we can ensure that the inserted loxP site will be highly permissive for expression of an inserted lentiviral genome construct.

Figure 5 shows the n-fold difference in GFP : β-actin ratio for clones EV1 A to D and EV2 A to D. All ratios are defined relative to a calibrator sample, defined as a ratio of 1.0. The calibrator sample used was RNA from EV2 cells transduced with Trap2 at an MOI of 0.3.

This identified the best expressing lines as:

- EV1 clone A - GFP : β-actin ratio is 22.8
- EV2 clone D - GFP : β-actin ratio is 18.6

These two lines were carried forward for further study.

**Excision of internal expression cassette by Cre recombinase**

The process of retroviral integration copies the loxP-containing modified 3' U3 to the 5' position. Therefore, one can excise the majority of the MLV Trap2 integration by the action of Cre recombinase. This will leave a single modified LTR, suitable for lentiviral genome integration.

EV1 clone A and EV2 clone D were transfected by the Cre expression plasmid pBS185 (Life Technologies). After one week, the cells were analysed for GFP by FACS (see Figure 13) to determine the excision frequency. This was measured at 20-70% in all lines.

Recombined clones were identified by limit dilute cloning cells and checking by microscope and FACS for loss of GFP expression.

**Insertion of EIAV genome into loxP site**

Plasmid pONY8x-loxP and pBS185 were co-transfected using Fugene into EV1 clone A (excised) and EV2 clone D (excised). A control transfection of pONY8z-loxP in the absence of pBS185 was also carried out.

Figure 14 shows lacZ expression of transfected cells withand without the addition of Cre recombinase (pBS185). The efficiency of the insertion event was estmated to be ~12% by computer image analysis.

We analysed cells for lacZ expression by FACS using the FluoReporter lacZ Flow Cytometry kit. The top 5% of lacZ positive cells were sorted by FACS and clones were derived by limiting dilution. In total, 12 clones of EV1/A/pONY8z-loxP were derived and 13 clones of EV2/D/pONY8z-loxP.

**EXAMPLE 7**

**Construction of EIAV vectors with LTR driven open reading frames**

The EIAV vector configurations described previously utilise a single promoter - transgene cassette located internally in the vector. For example in pONY8Z the promoter-transgene cassette is CMV-LacZ. However for some uses it would be advantageous to have the option of expressing a gene from the 5'LTR promoter as well. For example a marker gene such as green fluorescent protein (GFP), a resistance marker such as neomycin phosphotransferase (neo) or another protein or a biologically active entity such as a ribozyme. Previous experiments have shown that the EIAV LTR is weakly active in human cells in the absence of EIAV tat. However the transcriptional activity of the LTR can be increased by replacement of the EIAV U3 region with the MLV U3 region or the CMV promoter. This is achieved by introducing these alterations in the 3'LTRs of the vector plasmids. As a result of the replicative strategy of retroviruses the modified 3'LTR becomes positioned at the 5'end of the integrated vector and can thus drive expression of a gene placed downstream of the gag region. To ensure optimal levels of expression there should preferably be no ATG start codons prior to the start codon of the gene to be expressed. In pONY8Z the ATG start codon of gag and the next ATG downstream were mutated to ATTG in order to ablate expression of the

aminoterminal portion of gag present in the vectors, however there are 7 other ATG codons further downstream of these, within gag, from which translation might be initiated.

Described below are the replacement of the U3 region of EIAV with MLV or CMV promoters and the mutation of ATG codons in the gag region

**Replacement of the EIAV U3 region with MLV U3 or CMV promoters**

The MLV U3 region was introduced into pONY8Z vector by replacement of the 3'LTR with a synthetic MLV/EIAV LTR made by the overlapping PCR technique, using the following primers and templates.
The EIAV PPT/U3 sequence was amplified from pONY8.1Z using primers:

KM001: CAAAGCATGCCTGCAGGAATTCG (SEQ ID No 1)

and

KM003:
GCCAAACCTACAGGTGGGGTCTTTCATTATAAAACCCCTCATAAAAACCCCAC AG (SEQ ID No 2)

to give the following product:

CAAAGCATGCCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGAATTG GAAGAGCTTTAAATCCTGGCACATCTCATGTATCAATGCCTCAGTATGTTTAG AAAAACAAGGGGGGAACTGTGGGGTTTTTATGAGGGGTTTTATAATGAAAGA CCCCACCTGTAGGTTTGGC (SEQ ID No 3)

The MLV U3 region was amplified from pHIT111 (Soneoka et al., (1995) Nucleic Acids Res. **23**, 628-633) using KM004:

CTGTGGGGTTTTTATGAGGGGTTTTATAATGAAAGACCCCACCTGTAGGTTTG GC (SEQ ID No 4)

and

KM005:

GAAGGGACTCAGACCGCAGAATCTGAGTGCCCCCCGAGTGAGGGGTTGTGGG CTCT (SEQ ID No 5) to give the following product:

CTGTGGGGTTTTTATGAGGGGTTTTATAATGAAAGACCCCACCTGTAGGTTTGGCAAGCTAGCT
TAAGTAACGCCATTTTGCAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGTTCAGATC
AAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTT
CCTGCCCCGGCTCAGGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCT
GTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCAGC
CCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCC
TGTGCCTTATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCCG
AGCTCAATAAAAGAGCCCACAACCCCTCACTCGGGGGGCACTCAGATTCTGCGGTCTGAGTCC
CTTC (SEQ ID No 6)

The MLV U3/EIAV R/U5 was amplified from pONY8.1Z using primers

KM002: GAGCGCAGCGAGTCAGTGAGCGAG (SEQ ID No 7) and

KM006:

AGAGCCCACAACCCCTCACTCGGGGGGCACTCAGATTCTGCGGTCTGAGTCC CTTC (SEQ ID No 8)

to give the following product:

AGAGCCCACAACCCCTCACTCGGGGGGCACTCAGATTCTGCGGTCTGAGTCCCTTCTCTGCTG
GGCTGAAAAGGCCTTTGTAATAAATATAATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCTGTT
CGAGATCCTACAGAGCTCATGCCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTG
TTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGC
CTAATGAGTGAGCTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAAC

CTGTCGTGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGG
CGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC (SEQ ID No 9)

The PCR products described above were purified and then used as templates in new PCR reactions to link them together to obtain a 992bp product. The final product contains two SapI sites which flank the hybrid LTR. These allow introduction of the PCR product into the corresponding SapI sites present in the pONY8Z or pONY8.1Z vector plasmid, thereby creating pONY8Z MLVHyb and pONY8.1 MLVHyb. The sequence of the hybrid LTR in these plasmids was confirmed by sequencing. The titres obtained from the vectos in transient transfection assays are shown in Table 1. The titres were very similar to the titres from the parental construct pONY8Z and pONY8.1Z indicating that replacement of the EIAV U3 region with that of MLV had little or no detrimental effect on the infectious cycle of the vectors.

**Table 1. Titres obtained from MLV hybrid LTR vector plasmids**

| vector plasmid | #titre (l.f.u./ml) |
|---|---|
| pONY8Z | $3 \times 10^5$ |
| pONY8Z MLVHyb | $1 \times 10^5$ |
| pONY8.1Z | $6 \times 10^4$ |
| pONY8.1Z MLVHyb | $2 \times 10^4$ |

# Titre was measured on D17 cells and is expressed as LacZ forming units/ml (l.f.u./ml). Transfections were carried out in 293T cells using the vector plasmid shown and pRV67 (VSV-G expression plasmid), and pONY3.1 (EIAV gag/pol expression plasmid).

The structure of pONY8.1Z MLVHyb is shown in Figure 15 and the sequence of this plasmid is shown as SEQ ID No 10.

The EIAV promoter was also replaced by the human cytomegalovirus (CMV) promoter using a similar strategy. The primers and templates were the same except that KM003 was replaced by KM008:

GGCCATCGTGCCTCCCCACTCCTGCAGTTATAAAACCCCTCATAAAAACCCCA CAG (SEQ ID No 11)

KM004 was replaced by KM009:

CTGTGGGGTTTTTATGAGGGGTTTTATAAACTGCAGGAGTGGGGAGGCACGA TGGCC (SEQ ID No 12)

KM005 was replaced by KM010:

GAAGGGACTCAGACCGCAGAATCTGAGTGCCCGGTTCACTAAACGAGCTCTG CTTATATAGACC (SEQ ID No 13) and

KM006 was replaced by KM011:

GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGGGCACTCAGATTCTGCG GTCTGAGTCCCTTC (SEQ ID No 14)

The template for the PCR reaction with primers KM009 and KM010 was pONY2.1LacZ. This plasmid contains a single CMV promoter. The combined PCR product of 1319 bp was digested with SapI and introduced into the pONY8Z or pONY8.1Z backbone as described above for pONY8Z MLVHyb.

**Mutation of remaining ATG codons in the gag of pONY8Z to ATTG**

The alignment of the sequence of the leader and gag region present in vectors pONY4Z (an earlier generation EIAV vector), pONY8Z and a derivative of pONY8Z in which the 7 remaining ATG codons are mutated to ATTG is shown in Figure 16. These mutations were created by PCR mutagenesis as follows. The template for the PCR reactions was pONY8Z and the primers were:

F1: CGAGATCCTACAGTTGGCGCCCGAACAG (SEQ ID No 15);

R1:GAGTTACAATCTTCCAGCAATGGAATGACAATCCCTCAGCTGCCAGTCCTTTTCTTTTACAAAGTTGGTATCAATGAAATAAGTCTACTAGACTTAGC (SEQ ID No 16);

F2:TTCCATTGCTGGAAGATTGTAACTCAGACGCTGTCAGGACAAGAAAGAGAGGCCTTTGAAAGAACATTGGTGGGCAATTTCTGCTGTAAAGATTG (SEQ ID No 17);

R2:CAATATTTCGCTCTTAGGAGCTGGAATGATGCCTTTCCAATCTACTACAATTATTAATCTGGAGGCCCAATCTTTACAGCAGAAATTGCCCACCAATG (SEQ ID No 18);

R3:CCACTAGTTCTAGAGATATTCTTCAGAGGGCTCAGACTGCTTTTTATTAGCAGTCTTCTTTTCAATATTTCGCTCTTAGGAGC (SEQ ID No 19)

In the first stage of construction two PCR reactions were set up with primer pairs F1/R1 and F2/R2, respectively. These were purified and then used in a second 'overlapping' reaction in which primers F1 and R3 were added after 10 cycles. This procedure results in a 552bp PCR product (SEQ ID No 20):

CGAGATCCTACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAA
CCTGGCTGATCGTAGGATCCCCGGGACAGCAGAGGAGAACTTACAGAAGTCTTCTGGAGGTGT
TCCTGGCCAGAACACAGGAGGACAGGTAAGATTGGGAGACCCTTTGACATTGGAGCAAGGCG
CTCAAGAAGTTAGAGAAGGTGACGGTACAAGGGTCTCAGAAATTAACTACTGGTAACTGTAAT
TGGGCGCTAAGTCTAGTAGACTTATTTCATTGATACCAACTTTGTAAAAGAAAAGGACTGGCA
GCTGAGGGATTGTCATTCCATTGCTGGAAGATTGTAACTCAGACGCTGTCAGGACAAGAAAGA
GAGGCCTTTGAAAGAACATTGGTGGGCAATTTCTGCTGTAAAGATTGGGCCTCCAGATTAATA
ATTGTAGTAGATTGGAAAGGCATCATTCCAGCTCCTAAGAGCGAAATATTGAAAAGAAGACTG
CTAATAAAAAGCAGTCTGAGCCCTCTGAAGAATATCTCTAGAACTAGTGG

This was digested with NarI and XbaI and ligated into pONY8Z. pONY8Z MLVHyb and pONY8Z CMVHyb. which had been prepared for ligation by digestion with the same enzymes. These plasmids were designated pONY8ZA or pONY8ZA MLVHyb and pONY8ZA CMVHyb. The sequence for pONY8ZA CMVHyb is provided in SEQ ID No 52. These plasmids have a unique XbaI site into which can be inserted genes such as

GFP or neomycin phosphotransferase or other biologically active entity. This use of th site is demonstrated for GFP. The GFP ORF was obtained from pEGFP-1 (Clontech) by digestion with SmaI and XbaI, and then the ends filled in by treatment with T4 DNA polymerase. This fragment was then ligated into pONY8ZA or pONY8ZA MLVHyb and pONY8ZA CMVHyb prepared for ligation by digestion with XbaI and subsequent filling in with T4DNA polymerase. The resulting vector plasmids were called pONY8GZA or pONY8GZA MLVHyb and pONY8GZA CMVHyb. Other genes can be inserted at this site by manipulations apparent to those skilled in the art.

**Creation of EIAV vector genomes containing loxP sites in their LTR's**

The time taken to construct producer cell lines for EIAV vectors would be greatly reduced if it was possible to 1) locate and 2) reutilise a site in the host cell chromosome which was particularly favourable for high levels of transcription of the vector genome. In outline, this can be achieved by engineering loxP sites in the 3'LTR of EIAV vectors, transduction of the packaging cell line with vectors which carry loxP and hybrid LTRs, selection of cells which express the highest levels of vector genome and exchange of the test EIAV vector genome for the vector genome of choice using the cre/loxP recombination system.

The proposed scheme was evaluated using a derivative of pONY8GZA CMVHyb in which a loxP site was introduced into the PstI site between the EIAV sequences (required for efficient integration) and the CMV promoter in the 3'LTR. After transduction the integrated vector will thus have a loxP-CMV cassette located in the 5'LTR and 3'LTR's and therefore full length transcripts of the vector genome will be driven by the CMV promoter, which is a powerful promoter. pONY8GZA CMVHyb contains many PstI sites hence it was modified to allow insertion of the loxP site by digestion with XbaI and NheI and religation to create the subclone, pONY CMVHyb. This plasmid has a unique PstI site in the hybrid LTR. The loxP site was inserted into this site using two complimentary oligonucleotides which when annealed formed PstI-compatible termini. These were termed loxP POS

PSTI [GATAAC TTCGTATAATGTATGCTATACGAAGTTATCTGCA] (SEQ ID No 21)] and

loxP NEG PstI [GATAACTTCGTATAGCATACATTATACGAAGTTATCTGCA] (SEQ ID No 22)

The sequence and orientation of the loxP site was confirmed by DNA sequencing and the plasmid called pONY CMVHyb loxP. The central part of the vector genome was then reintroduced into this subclone by transfer of the NotI-BstEII fragment from pONY8GZA CMVHyb into pONY CMVHyb cut the same way. The resulting vector was termed pONY8GZA CMVHyb loxP.

Two routes for construction of a producer cell line are available using pONY8GZA CMVHyb loxP. The plasmid can be introduced into a packaging cell line by transfection or vector particles can be made using the 293T and these used to transduce the packaging cell line. Since the vector is derived from EIAV, rather than MLV, it is able to transduce non-dividing cells or slowly dividing cells. In this situation it has been hypothesised that integrations occur at chromosomal sites that are constitutively open; that is, are likely to be sites at which high levels of transcription will be maintained for extended periods. This may be important for the long term usefulness of the producer cell line and thus represents an advantage of strategy using transduction.

Producer cell lines were made by transfection or transduction of a TE671-derived cell line (EV11E) which has stably integrated copies of VSV-G and the synthetic EIAV gag/pol under the control of CMV promoters. Prior to transfection with pONY8GZA CMVHyb loxP it was linearised by digestion with AhdI. Seven days following transfection or transduction cells expressing the highest levels of GFP were selected by FACS and then cloned by limiting dilution. A number of clones were analysed for levels of full length vector RNA using Taqman technology based assays in order to confirm the hypothesis that the highest level of GFP expression correlates with the highest levels of vector RNA.

The cell line which expressed the highest level of RNA was then tested for production of transducing vector particles 5 days after changing the temperature of incubation from 37C

to 32°C. At 32°C. VSV-G expression is induced however maximal levels of VSV-G are only obtained after 5 days at the permissive temperature (see WO 00/52188). The cell line producing the highest titre, EV11E CMV loxP was selected for further work.

In order to exchange the vector genome with for another EIAV vector genome with a more suitable configuration for use in the clinical setting EV11EloxP cells were transfected with cre recombinase expression plasmid, pBS185 (Gibco), which results in excision of the EIAV vector between the loxP sites. This leaves a loxP-CMV promoter R-U5 sequence in the cells. Cells from which the EIAV vector genome had been excised were selected on the basis of low levels of GFP expression by FACS and assumed to be clonal on the basis of the clonality of EV11E CMV loxP. These were termed EV11EloxPΔ and used as targets for new EIAV vector genomes.

**Construction of EIAV vector genomes with downstream REV expression cassettes and flanking loxP sites**

The production of vector particles from minimal EIAV vectors (those which do not express EIAV REV or any other EIAV proteins) is increased by about 10-fold in the presence of EIAV REV in our 293T transient production system when the codon-optimsed EIAV gag/pol expression construct is used to drive production of vector particles as set out in Table 2. This may be improved nuclear to cytoplasmic transport of the vector genome in the presence of REV protein. Packaging/producer cell lines for EIAV vectors may be engineered to express Rev protein. One approach would be to engineer cells to express EIAV REV from an independently transfected expression cassette. However, the cassette and the vector genome may be subject to differential regulation, for example by methylation or chromosome remodelling. Such an effect may limit the useful life of such cell lines.

Table 2. Effect of REV expression on titres obtained from REV-expressing [REV+] and non-expressing [REV-] vectors

| vector plasmid | gag/pol expression | [a]titre |
|---|---|---|

| | plasmids | (l.f.u./ml) |
|---|---|---|
| pONY4Z [REV+] | pONY3.1 | 2.0±0.4 x $10^6$ |
| pONY4Z [REV+] | pE SYN GP | 0.9±0.2 x $10^6$ |
| pONY8Z [REV-] | pONY3.1 | 1.5±0.2 x $10^6$ |
| pONY8Z [REV-] | pE SYN GP | 1.9±0.6 x $10^5$ |

[#] Titre was measured on D17 cells and is expressed as LacZ forming units/ml (l.f.u./ml). Transfections were carried out in 293T cells using the vector plasmid and gag/pol expression plasmid shown and pRV67 (VSV-G expression plasmid) (See WO 00/52188).

REV+ and REV- reflect the rev expression status of the vectors. REV+ reflects vectors which express the REV protein. REV- reflects expression vectors which do not express the REV protein.

**pESYNGP**

The gag/pol expression plasmid shown in Figure called called pESYNGP was constructed as follows: The codon-optimised EIAV gag/pol ORF was synthesised by Operon Technologies Inc., Alameda and supplied in a proprietary plasmid backbone, GeneOp. The complete fragment synthesised included sequences flanking the EIAV gag/pol ORF: tctagaGAATTCGCCACC**ATG- EIAV gag/pol- UGA**ACCCGGGgcggccgc (SEQ ID No 44). The ATG start and UGA stop codons are shown in bold. XbaI and NotI sites are in lower case. These were used to transfer the gag/pol ORF from GeneOp into pCIneo (Promega) using the NheI and NotI sites in the latter.

**pESDSYNGP**

An alternative expression plasmid for expression of the synthetic EIAV gag/pol could also be used. It is called pESDSYNGP and its construction is described as follows: ESDSYNGP was made from pESYNGP by exchange of the 306bp EcoRI-NheI fragment, from just upstream of the start codon for gag/pol to approximately 300 base pairs inside

the gag/pol ORF with a 308bp EcoRI-NheI fragment derived by digestion of a PCR made using pESYNGP as template and using the following primers: SD FOR [GGCTAGAGAATTCCAGGTAAGATGGGCGATCCCCTCACCTGG] (SEQ ID No 60) and SD REV [TTGGGTACTCCTCGCTAGGTTC] (SEQ ID No 61). This manipulation replaces the Kozak concensus sequence upstream of the ATG in pESYNGP with the splice donor found in EIAV. The sequence between the EcoRI site and the ATG of gag/pol is thus CAGGTAAG (SEQ ID No 62).

The sequences for pESYNGP (SEQ ID No 53) and pESDSYNGP (SEQ ID No 54) are provided.

Packaging/Producer cells may be engineered by physically linking the genome and EIAV REV expression cassettes. In this way stable transfectants may be generated which contain the vector genome and the EIAV REV expression cassette in the same chromatin environment. This manipulation may ensure that the relative levels of transcription of the vector genome and the REV expression cassette are maintained leading to an increased duration of vector production from the producer cells.

Previous work has suggested that optimisation of the level of REV may be required with respect to the level of vector genome (see WO 98/17815). We have examined the levels of vector production in a transient system in which several different promoters are used to drive REV expression in order to determine which vector genome-rev expression cassette is optimal for use in constructing producer cell lines. The highest titres were obtained with FB29 and PGK promoters driving REV expression.

The following describes the construction of EIAV vector genomes plasmids in which there is a downstream expression cassette for synthetic EIAV REV protein. The promoters tested were FB29, PGK, TK, CMV, SV40 and RSV. In addition the loxP sites were engineered into the vector plasmid backbone in such a way that the genome and introduced promoter-REV expression plasmid was flanked. In this way, the complete vector-REV cassette can be recombined into loxP sites in the target cell.

The complete construction of the FB29 and PGK containing plasmids is described here. The REV expression construct was inserted in the both orientations with respect to the EIAV vector genome. Plasmids in which the FB29 or PGK promoters drive REV expression are being utilised for construction of stable producer cell lines.

**Construction of plasmids**

In the first step of construction an SfiI site was inserted downstream of the EIAV vector sequence. This site is the insertion site for the promoter REV cassettes. The construction was made as follows. pONY8Z was digested with EheI and NruI, the ends were blunted by treatment with T4 DNA polymerase and religated. The resulting plasmid, pONY8Z delta, is thus deleted with respect to the leader, gag, reporter cassette and most of the Rev/RRE regions.

pONY8Z delta was mutated to contain loxP sites inserted in the DraII site immediately to the 5'of the CMV promoter and in the BspLU11I site to the 3' of the vector genome. The loxP sites were inserted using complementary nucleotide pairs which when annealed had overhanging termini suitable for cloning into these sites and were inserted in two steps of cloning. The oligonucleotides for insertion into the DraIII site were

VSAT 158: [GTGATAACTTCGTATAATGTATGCTATACGAAGTTATCACTAC] (SEQ ID No 23)

and

VSAT 155 [GTGATAACTTCGTATAGCATACATTATACGAAGTTATCACGTA] (SEQ ID No 24)

For the BspLU11I they were:

VSAT 156 [CATGTATAACTTCGTATAATGTATGCTATACGAAGTTATA] (SEQ ID No 25) and

VSAT 157 [CATGTATAACTTCGTATAGCATACATTATACGAAGTTATA] (SEQ ID No 26)

Plasmids were selected in which the orientation of the loxP at both sites were the same and the same as the EIAV vector genome. The modified plasmid was called pONY8Z delta 2xloxP.

pONY8Z delta 2xloxP has a unique PvuII site downstream of the deleted EIAV vector genome into which annealed complementary oligonucleotides encoding SfiI sites were inserted. The oligonucleotides were:

SFI SRFPOS [AGTAGGCCGCCTCGGCCGCCCGGGCATCA] (SEQ ID No 27) and

SFI SRF NEG [TGATGCCCGGGCGGCCGAGGCGGCCTACT] (SEQ ID No 28)

Clones which had the SfiI – SrfI sites in either orientation were selected for further work. These were called pONY8Z delta SfiI FOR and REV.

**Amplification and cloning of FB29 and PGK promoters**

The FB29 promoter was amplified from pRDF (Cosset FL, et al.*J Virol* 1995 Dec;69(12):7430-6) using primers:

FB29 POS [TAGCCGAGATCTCAAATTGCTTAGCCTGATAGCC] (SEQ ID No 29)and

FB29 NEG [TGCGTAGCTAGCCTCCCGGTGGTGGGTCGGTG] (SEQ ID No 30) which introduce

5'BglII and 3'NheI sites.

The PGK promoter was amplified from pPE327 using primers

PGK POS [AGCAGTAGATCTGGGGTTGGGGTTGCGCCTTT] (SEQ ID No 31) and

PGK NEG [CGTCATGCTAGCCTGGGGAGAGAGGTCGGTG] (SEQ ID No 32)

The PGK promoter sequence obtained from this plasmid was the same as the sequence of GenBank Acc. No. M11958 except that it has a single mutation: nucleotide 347 of M11958 is changed from G to A. The TK promoter and intron was amplifed from pRL-TK (Promega) with:

TK POS [TACGGAAGATCTAAATGAGTCTTCGGACCT] (SEQ ID No 33) and

TK NEG [CTCAACGCTAGCGTACTCTAGCCTTAAGAGCTG] (SEQ ID No 34)

The RSV promoter was amplifed from pREP7 (Invitrogen) with

RSV POS [TACCAGAGATCTTCTAGAGTCGACCAATTCTCATG] (SEQ ID No 35) and

RSV NEG [CATCGAGCTAGCAGCTTGGAGGTGCACACCAATG] (SEQ ID No 36) and

The SV40 promoter was amplifed from pCIneo (Promega) with:

SV40 POS [GATGGTAGATCTGCGCAGCACCATGGCCTGAA] (SEQ ID No 37) and

SV40 NEG [CTCGAAGCTAGCAGCTTTTTGCAAAAGCCTAGGC] (SEQ ID No 38)

The PCR fragments were digested with BglII and NheI and ligated into pSL1180 (Pharmacia) which had been prepared by digestion using the same enzymes. Following transformation into E.coli DNA was prepared and the sequence of the promoters checked by DNA sequencing. Clones in which the correct promoter sequence was present were used for further work and were called pSL1180-FB29, pSL1180-PGK, pSL1180-RSV, pSL1180-SV40, pSL1180-TK.

In the next step the promoter fragments were positioned to drive transcription of synthetic EIAV REV in pE syn REV. pE syn REV is a pCIneo based plasmid (Promega) which was made by introducing the EcoRI to SalI fragment from the synthetic EIAV REV plasmid into the polylinker region of the plasmid using the same sites. The synthetic

EIAV REV plasmid made by Operon contains a codon-optimised EIAV REV open reading frame flanked by EcoRI and SalI. The sequence of this fragment is shown as SEQ ID No 39.

Prior to replacement of the CMV promoter in pE syn REV it was modified as follows. The SV40 neo region was deleted by digestion with KpnI and BamHI. the ends blunted by treatment with T4 DNA polymerase and then religated. The plasmid was termed pE syn REV delta. Next SfiI sites were introduced into both the BglII site which is just 5' of the CMV promoter and DraIII site downstream of the polyA signal.

The oligonucleotides used for this were as follows:

SFI FOR BglII POS [GATCGGCCGCCTCGGCCA] (SEQ ID No 40) and

SFI FOR BglII NEG [ GATCTGGCCGAGGCGGCC] (SEQ ID No 41)and

SFI FOR DRA POS [ GGCCGCCTCGGCCGTA] (SEQ ID No 42) and

SFI FOR DRA NEG [GGCCGAGGCGGCCTAC] (SEQ ID No 43)

Clones in which the SfiI was located 5' of the BglII site were selected were used for further work. The plasmid obtained after this two step manipulation was termed pE syn REV delta 2xSfiI. It has the following features: 5'SfiI sites – BglII site - CMV promoter and intron – NheI site – E syn REV - polyA site – 3'SfiI site. Hence the CMV promoter can be excised by digestion with BglII and NheI and replaced with the promoter of choice obtained from the pSL1180 series of clones by digestion with the same enzymes. Construction details are included from this point for only the constructs which contained FB29 and PGK promoters, however a similar scheme was used for the other promoters, except that a partial SfiI digestion was required for transfer of the SV40-REV cassette.

Promoter fragment were obtained from pSL1180 – FB29 and pSL1180 – PGK by digestion with BglII and NheI and ligated into pE syn REV delta 2xSfiI digested with the

same enzymes. The resulting plasmids were called FB29 E SYN REV and PGK E SYN REV.

In the next stage the internal regions of pONY8G, pONY8.1G SIN MIN and pONY4G were obtained by digestion with SgfI (which has unique site in the CMV promoter driving the EIAV vector genome) and MunI (which cuts in the 3'LTR) and ligated in to pONY8Z delta SfiI FOR and REV prepared for ligation by digestion with the same enzymes. The resulting plasmids were called pONY8G SfiI FOR and REV, pONY8.1G SIN MIN SfiI FOR and REV and pONY4G SfiI FOR and REV.

In the final stage the promoter-REV cassettes were moved from FB29 E SYN REV and PGK E SYN REV into pONY8G SfiI FOR and REV, pONY8.1G SIN MIN SfiI FOR and REV and pONY4G SfiI FOR. This manipulation was achieved as follows. FB29 E syn REV, PGK E syn REV, and the vector plasmids described immediately above were digested with SfiI and ligations set up with appropriates fragments. The promoter-REV cassettes were orientated in the same or opposite orientations with respect to the EIAV vector genome in the 'FOR' and 'REV' plasmids The resulting plasmids were called pONY8.3G FB29 + or – and pONY8.3G PGK+ and –. A schematic structure of the pONY 8.3 +/- plasmids is shown in Figure 17.

The performance of these constructs was tested in relation to pONY8G in 293T transient production assays and the results are shown in Table 3.

**The sequence of the *EcoR*I to *Sal*I fragment representing the codon-optimised EIAV REV open reading frame obtained from the plasmid synthesised by Operon (SEQ ID No 39)**

EcoRI and SalI sites are in bold. The ATG start and UGA termination codons are underlined

GAATTCGCCACC<u>ATG</u>GCTGAGAGCAAGGAGGCCAGGGATCAAGAGATG

ACCTCAAGGAA
GAGAGCAAAGAGGAGAAGCGCCGCAACGACTGGTGGAAGATCGACCCA
AAGGCCCCCTG
GAGGGGGACCAGTGGTGCCGCGTGCTGAGACAGTCCCTGCCCGAGGAGAAGATTCCT
AGC
CAGACCTGCATCGCCAGAAGACACCTCGGCCCCGGTCCCACCCAGCACACACCCTCC
AGA
AGGGATAGGTGGATTAGGGGCCAGATTTTGCAAGCCGAGGTCCTCCAAGAAAGGCTG
GAA
TGGAGAATTAGGGGCGTGCAACAAGCCGCTAAAGAGCTGGGAGAGGTGAATCGCGG
CATC
TGGAGGGAGCTCTACTTCCGCGAGGACCAGAGGGGCGATTTCTCCGCATGGGGAGGC
TAC
CAGAGGGCACAAGAAAGGCTGTGGGGCGAGCAGAGCAGCCCCCGCGTCTTGAGGCC
CGGA
GACTCCAAAAGACGCCGCAAACACCTG<u>TGA</u>**AGTCGAC**

Table 3

Titres obtained from a representative experiment in which the vector-REV constructs were tested by transient 293T production assay. The vector constructs were cotransfected with pE synGP, the synthetic EIAV gag/pol expression plasmid, and pRV67, VSV-G expression plasmid. Titres were measured in D17 cells.

| <u>Plasmid</u> | Titre (g.f.u./ml) |
|---|---|
| pONY8G SfiI FOR | $1.6\times10^4$ |
| #pONY8G SfiI FOR PLUS pE syn REV | $5.2\times10^5$ |
| pONY8.3G FB29 + | $8.8\times10^5$ |
| pONY8.3G FB29 - | $7.8\times10^5$ |
| pONY8.3G PGK + | $1.2\times10^6$ |
| pONY8.3G PGK - | $1.2\times10^6$ |
| pONY8G | $9.4\times10^5$ |

Titre was assessed on D17 cells and is expressed as green fluorescent protein cell units/ml (g.f.u./ml). Transfections were carried out with pE syn GP KOZAK and pRV67 as described previously.

* pONY8G SfiI FOR is identical to the pONY8.3 derivatives except that there is no promoter-REV expression cassette is inserted in the SfiI site

# pE syn REV plasmid was also included in this transfection

pONY8G is a standard EIAV vector genome used for comparative purposes

pONY8.3G FB29 – is shown as SEQ ID No 45

pONY8.3G FB29 + is shown as SEQ ID No 46

pONY8.3GPGK – is shown as SEQ ID No 47

pONY8.3G PGK + is shown as SEQ ID No 48.

## SUMMARY

Thus, in summation, the present invention provides high titer regulated retroviral vectors. These regulated retroviral vectors include lentivectors, HRE-regulated vectors and functional SIN vectors which can be produced at high titres from derived producer cell lines.

The present invention also provides a method other than retroviral transduction for the transfer of a regulated retroviral vector into a derived producer cell line. This method comprises a recombinase assisted method which allows for the production of high titer regulated retroviral vectors.

In one broad aspect, the present invention relates to the selection of cells which express high levels of a retroviral vector genome and exchange of this retroviral genome for the

vector genome of choice, preferably a regulated retroviral vector genome or a lentiviral vector genome using a cre/loxP recombination system. Thus, the present invention enables regulated retroviral vectors to be produced at high titres from transduced producer cell lines.

In another broad aspect, the present invention relates to the selection of cells which express high levels of a retroviral vector genome and exchange of this retroviral genome for the vector genome of choice, preferably a regulated retroviral vector genome or a lentiviral vector genome using a cre/loxP recombination system and a retroviral vector producton system which incorporates a REV protein production system. Thus, the present invention enables regulated retroviral vectors to be produced at high titres from transduced producer cell lines.

All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the described methods and system of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

# SEQUENCE LISTING PART OF THE DESCRIPTION

## pONY8.1Z MLVHyb (SEQ ID NO 10)

AGATCTTGAATAATAAAATGTGTGTTTGTCCGAAATACGCGTTTTGAGATTTCTGTCGCCGACTAAATTCATGTCGCGCG
ATAGTGGTGTTTATCGCCGATAGAGATGGCGATATTGGAAAAATTGATATTTGAAAATATGGCATATTGAAAATGTCGC
CGATGTGAGTTTCTGTGTAACTGATATCGCCATTTTTCCAAAAGTGATTTTTGGGCATACGCGATATCTGGCGATAGCGC
TTATATCGTTTACGGGGGATGGCGATAGACGACTTTGGTGACTTGGGCGATTCTGTGTGTCGCAAATATCGCAGTTTCGA
TATAGGTGACAGACGATATGAGGCTATATCGCCGATAGAGGCGACATCAAGCTGGCACATGGCCAATGCATATCGATC
TATACATTGAATCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATT
GCATACGTTGTATCCATATCGTAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCATGTTGACATTGATTATT
GACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGG
TAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCC
AATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCAT
ATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGG
ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG
CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA
ATCAACGGGACTTTCCAAAATGTCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGG
TGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGGGCACTCAGATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCT
GAAAAGGCCTTTGTAATAAATATAATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATCCTACAGTTGGCGC
CCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGGCTGATCGTAGGATCCCCGGGACAGCAGAGGA
GAACTTACAGAAGTCTTCTGGAGGTGTTCCTGGCCAGAACACAGGAGGACAGGTAAGATTGGGAGACCCTTTGACATT
GGAGCAAGGCGCTCAAGAAGTTAGAGAAGGTGACGGTACAAGGGTCTCAGAAATTAACTACTGGTAACTGTAATTGGG
CGCTAAGTCTAGTAGACTTATTTCATGATACCAACTTTGTAAAAGAAAAGGACTGGCAGCTGAGGGATGTCATTCCATT
GCTGGAAGATGTAACTCAGACGCTGTCAGGACAAGAAAGAGAGGCCTTTGAAAGAACATGGTGGGCAATTTCTGCTGT
AAAGATGGGCCTCCAGATTAATAATGTAGTAGATGGAAAGGCATCATTCCAGCTCCTAAGAGCGAAATATGAAAAGAA
GACTGCTAATAAAAAGCAGTCTGAGCCCTCTGAAGAATATCTCTAGAACTAGTGGATCCCCCGGGCTGCAGGAGTGGG
GAGGCACGATGGCCGCTTTGGTCGAGGCGGATCCGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATA
TTGGCTATTGGCCATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCAT
GTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGC
GTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATG
TTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAG
TACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTG
GCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGT
TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGC
ATGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTT
TGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGCCCCAAGCTTCAGCTGCTCGAGGATCTGCGGATCCGG
GGAATTCCCCAGTCTCAGGATCCACCATGGGGGATCCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTAC
CCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCC
CAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGTGCCGGAAAGCTGGCTG
GAGTGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGATGCGCCCATCTACA
CCAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGTTACTCGCTCACATTT
AATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGT
GCAACGGGCGCTGGGTCGGTTACGGCCAGGACAGTCGTTTGCCGTCTGAATTTGACCTGAGCGCATTTTTACGCGCCGG
AGAAAACCGCCTCGCGGTGATGGTGCTGCGTTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATGTGGCGGATGAG
CGGCATTTTCCGTGACGTCTCGTTGCTGCATAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTTAATG
ATGATTTCAGCCGCGCTGTACTGGAGGCTGAAGTTCAGATGTGCGGCGAGTTGCGTGACTACCTACGGGTAACAGTTTC
TTTATGGCAGGGTGAAACGCAGGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGAGCGTGGTGGTTAT

GCCGATCGCGTCACACTACGTCTGAACGTCGAAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCTCTATCGTGCGG
TGGTTGAACTGCACACCGCCGACGGCACGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCGGATTGA
AAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAACCGTCACGAGCATCATCCTCTGCATGGT
CAGGTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAACAACTTTAACGCCGTGCGCTGTTCGC
ATTATCCGAACCATCCGCTGTGGTACACGCTGTGCGACCGCTACGGCCTGTATGTGGTGGATGAAGCCAATATTGAAAC
CCACGGCATGGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGATGAGCGAACGCGTAACGCGAAT
GGTGCAGCGCGATCGTAATCACCCGAGTGTGATCATCTGGTCGCTGGGGAATGAATCAGGCCACGGCGCTAATCACGA
CGCGCTGTATCGCTGGATCAAATCTGTCGATCCTTCCCGCCCGGTGCAGTATGAAGGCGGCGGAGCCGACACCACGGCC
ACCGATATTATTTGCCCGATGTACGCGCGCGTGGATGAAGACCAGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAA
AATGGCTTTCGCTACCTGGAGAGACGCGCCCGCTGATCCTTTGCGAATACGCCCACGCGATGGGTAACAGTCTTGGCGG
TTTCGCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTACAGGGCGGCTTCGTCTGGGACTGGGTGGATCAGTCGC
TGATTAAATATGATGAAAACGGCAACCCGTGGTCGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTT
CTGTATGAACGGTCTGGTCTTTGCCGACCGCACGCCGCATCCAGCGCTGACGGAAGCAAAACACCAGCAGCAGTTTTTC
CAGTTCCGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGAGCTCCTGCACT
GGATGGTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAGGTAAACAGTTGA
TTGAACTGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGGCTCACAGTACGCGTAGTGCAACCGAACGCGA
CCGCATGGTCAGAAGCCGGGCACATCAGCGCCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCG
CCGCGTCCCACGCCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCGAGCTGGGTAATAAGCGTTGGCAATT
TAACCGCCAGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACTGCTGACGCCGCTGCGCGATCAGTTC
ACCCGTGCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGCATTGACCCTAACGCCTGGGTCGAACGCTGG
AAGGCGGCGGGCCATTACCAGGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATT
ACGACCGCTCACGCGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGGAAAACCTACCGGATTGATGGTAGTGGTC
AAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCGGCGCGGATTGGCCTGAACTGCCAGCTGGC
GCAGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCGCAAGAAAACTATCCCGACCGCCTTACTGCCGCCTGTTTT
GACCGCTGGGATCTGCCATTGTCAGACATGTATACCCCGTACGTCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGC
GCGAATTGAATTATGGCCCACACCAGTGGCGCGGCGACTTCCAGTTCAACATCAGCCGCTACAGTCAACAGCAACTGAT
GGAAACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACATGGCTGAATATCGACGGTTTCCATATGGGGATTGG
TGGCGACGACTCCTGGAGCCCGTCAGTATCGGCGGAATTCCAGCTGAGCGCCGGTCGCTACCATTACCAGTTGGTCTGG
TGTCAAAAATAATAATAACCGGGCAGGGGGGATCCGCAGATCCGGCTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAG
TCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGAA
TTGGAAGAGCTTTAAATCCTGGCACATCTCATGTATCAATGCCTCAGTATGTTTAGAAAAACAAGGGGGGAACTGTGGG
GTTTTTATGAGGGGTTTTATAATGAAAGACCCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTGCAAGG
CATGGAAAAATACATAACTGAGAATAGAGAAGTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCA
AACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAAC
AGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGC
AGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCCTGTGCCTTATTTGAACTAACCAA
TCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAAGAGCCCACAACCCCTCACTCGGGGG
CACTCAGATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGGCCTTTGTAATAAATATAATTCTCTACTCAGTCC
CTGTCTCTAGTTTGTCTGTTCGAGATCCTACAGAGCTCATGCCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAA
ATTGTTATCCGCTCACAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAG
CTAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAATCG
GCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTT
CGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAG
AACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGC
CCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCG
TTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCG
GGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGT
GCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGAC
TTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGT

GGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAG
AGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGC
AGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAG
GGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAA
AGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCG
TTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCA
ATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGA
AGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAA
TAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCG
GTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTT
GTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGT
AAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCC
CGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCG
AAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTT
TTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGA
AATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATAT
TTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTAAATTGTAAGCGTT
AATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCC
TTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTG
GACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTT
TGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCAA
CCTGGCTTATCGAAATTAATACGACTCACTATAGGGAGACCGGC

## pONY8.3G FB29 – (SEQ ID No 45)

AGATCTTGAATAATAAAATGTGTGTTTGTCCGAAATACGCGTTTTGAGATTTCTGTCGCC
GACTAAATTCATGTCGCGCGATAGTGGTGTTTATCGCCGATAGAGATGGCGATATTGGAA
AAATTGATATTTGAAAATATGGCATATTGAAAATGTCGCCGATGTGAGTTTCTGTGTAAC
TGATATCGCCATTTTTCCAAAAGTGATTTTTGGGCATACGCGATATCTGGCGATAGCGCT
TATATCGTTTACGGGGGATGGCGATAGACGACTTTGGTGACTTGGGCGATTCTGTGTGTC
GCAAATATCGCAGTTTCGATATAGGTGACAGACGATATGAGGCTATATCGCCGATAGAGG
CGACATCAAGCTGGCACATGGCCAATGCATATCGATCTATACATTGAATCAATATTGGCC
ATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCCATATCGTAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCC
ATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
TAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAAT
AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGG
ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT
GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTGCGATCGCCCGCC
CCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGT
TTAGTGAACCGGGCACTCAGATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGG
CCTTTGTAATAAATATAATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATC
CTACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGG
CTGATCGTAGGATCCCCGGGACAGCAGAGGAGAACTTACAGAAGTCTTCTGGAGGTGTTC

CTGGCCAGAACACAGGAGGACAGGTAAGATTGGGAGACCCTTTGACATTGGAGCAAGGCG
CTCAAGAAGTTAGAGAAGGTGACGGTACAAGGGTCTCAGAAATTAACTACTGGTAACTGT
AATTGGGCGCTAAGTCTAGTAGACTTATTTCATGATACCAACTTTGTAAAAGAAAAGGAC
TGGCAGCTGAGGGATGTCATTCCATTGCTGGAAGATGTAACTCAGACGCTGTCAGGACAA
GAAAGAGAGGCCTTTGAAAGAACATGGTGGGCAATTTCTGCTGTAAAGATGGGCCTCCAG
ATTAATAATGTAGTAGATGGAAAGGCATCATTCCAGCTCCTAAGAGCGAAATATGAAAAG
AAGACTGCTAATAAAAAGCAGTCTGAGCCCTCTGAAGAATATCTCTAGAACTAGTGGATC
CCCCGGGCTGCAGGAGTGGGGAGGCACGATGGCCGCTTTGGTCGAGGCGGATCCGGCCAT
TAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATA
CGTTGTATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCAT
GTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATA
GCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC
CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACG
TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGAT
AGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT
TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGC
AAATGGGCGGTAGGCATGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACC
GTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACC
GATCCAGCCTCCGCGGCCCCAAGCTTGTTGGGATCCACCGGTCGCCACCATGGTGAGCAA
GGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAA
CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGAC
CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCAC
CCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTT
CTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA
CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT
CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTA
CAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGT
GAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCA
GCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCAC
CCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCGA
CTCTAGAGTCGACCTGCAGGCATGCAAGCTTCAGCTGCTCGAGGGGGGGCCCGGTACCCA
GCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGGGAAGTATTTATCACTAATCAAGCAC
AAGTAATACATGAGAAACTTTTACTACAGCAAGCACAATCCTCCAAAAAATTTTGTTTTT
ACAAAATCCCTGGTGAACATGATTGGAAGGGACCTACTAGGGTGCTGTGGAAGGGTGATG
GTGCAGTAGTAGTTAATGATGAAGGAAAGGGAATAATTGCTGTACCATTAACCAGGACTA
AGTTACTAATAAAACCAAATTGAGTATTGTTGCAGGAAGCAAGACCCAACTACCATTGTC
AGCTGTGTTTCCTGACCTCAATATTTGTTATAAGGTTTGATATGAATCCCAGGGGGAATC
TCAACCCCTATTACCCAACAGTCAGAAAAATCTAAGTGTGAGGAGAACACAATGTTTCAA
CCTTATTGTTATAATAATGACAGTAAGAACAGCATGGCAGAATCGAAGGAAGCAAGAGAC
CAAGAATGAACCTGAAAGAAGAATCTAAAGAAGAAAAAAGAAGAAATGACTGGTGGAAAA
TAGGTATGTTTCTGTTATGCTTAGCAGGAACTACTGGAGGAATACTTTGGTGGTATGAAG
GACTCCCACAGCAACATTATATAGGGTTGGTGGCGATAGGGGGAAGATTAAACGGATCTG
GCCAATCAAATGCTATAGAATGCTGGGGTTCCTTCCCGGGGTGTAGACCATTTCAAAATT
ACTTCAGTTATGAGACCAATAGAAGCATGCATATGGATAATAATACTGCTACATTATTAG
AAGCTTTAACCAATATAACTGCTCTATAAATAACAAAACAGAATTAGAAACATGGAAGTT

AGTAAAGACTTCTGGCATAACTCCTTTACCTATTTCTTCTGAAGCTAACACTGGACTAAT
TAGACATAAGAGAGATTTTGGTATAAGTGCAATAGTGGCAGCTATTGTAGCCGCTACTGC
TATTGCTGCTAGCGCTACTATGTCTTATGTTGCTCTAACTGAGGTTAACAAAATAATGGA
AGTACAAAATCATACTTTTGAGGTAGAAAATAGTACTCTAAATGGTATGGATTTAATAGA
ACGACAAATAAAGATATTATATGCTATGATTCTTCAAACACATGCAGATGTTCAACTGTT
AAAGGAAAGACAACAGGTAGAGGAGACATTTAATTTAATTGGATGTATAGAAAGAACACA
TGTATTTTGTCATACTGGTCATCCCTGGAATATGTCATGGGGACATTTAAATGAGTCAAC
ACAATGGGATGACTGGGTAAGCAAAATGGAAGATTTAAATCAAGAGATACTAACTACACT
TCATGGAGCCAGGAACAATTTGGCACAATCCATGATAACATTCAATACACCAGATAGTAT
AGCTCAATTTGGAAAAGACCTTTGGAGTCATATTGGAAATTGGATTCCTGGATTGGGAGC
TTCCATTATAAAATATATAGTGATGTTTTTGCTTATTTATTTGTTACTAACCTCTTCGCC
TAAGATCCTCAGGGCCCTCTGGAAGGTGACCAGTGGTGCAGGGTCCTCCGGCAGTCGTTA
CCTGAAGAAAAAATTCCATCACAAACATGCATCGCGAGAAGACACCTGGGACCAGGCCCA
ACACAACATACACCTAGCAGGCGTGACCGGTGGATCAGGGGACAAATACTACAAGCAGAA
GTACTCCAGGAACGACTGGAATGGAGAATCAGAGGAGTACAACAGGCGGCCAAAGAGCTG
GGTGAAGTCAATCGAGGCATTTGGAGAGAGCTATATTTCCGAGAAGACCAAAGGGGAGAT
TTCTCAGCCTGGGGCGGCTATCAACGAGCACAAGAACGGCTCTGGGGGGAACAATCCTCA
CCAAGGGTCCTTAGACCTGGAGATTCGAAGCGAAGGAGGAAACATTTATGACTGTTGCAT
TAAAGCCCAAGAAGGAACTCTCGCTATCCCTTGCTGTGGATTTCCCTTATGGCTATTTTG
GGGACTAGTAATTATAGTAGGACGCATAGCAGGCTATGGATTACGTGGACTCGCTGTTAT
AATAAGGATTTGTATTAGAGGCTTAAATTTGATATTTGAAATAATCAGAAAAATGCTTGA
TTATATTGGAAGAGCTTTAAATCCTGGCACATCTCATGTATCAATGCCTCAGTATGTTTA
GAAAAACAAGGGGGGAACTGTGGGGTTTTTATGAGGGGTTTTATAAATGATTATAAGAGT
AAAAAGAAAGTTGCTGATGCTCTCATAACCTTGTATAACCCAAAGGACTAGCTCATGTTG
CTAGGCAACTAAACCGCAATAACCGCATTTGTGACGCGAGTTCCCCATTGGTGACGCGTT
AACTTCCTGTTTTTACAGTATATAAGTGCTTGTATTCTGACAATTGGGCACTCAGATTCT
GCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGGCCTTTGTAATAAATATAATTCTCTA
CTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATCCTACAGAGCTCATGCCTTGGCGTAA
TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA
CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA
ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGTGATGCCCG
GGCGGCCGAGGCGGCCTACGTGAACCATCACCCAAATCAAGTTTTTTGCGGTCGAGGTGC
CGTAAAGCTCTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAG
CCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTG
GCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTA
CAGGGCGCGTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGG
CCTCTTCGCTATTACGCCAGCCCGGATCGATCCTTATCGGATTTTACCACATTTGTAGAG
GTTTTACTTGCTTTAAAAAACCTCCCACATCTCCCCCTGAACCTGAAACATAAAATGAAT
GCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGC
ATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAA
CTCATCAATGTATCTTATCATGTCTGCTCGAAGCATTAACCCTCACTAAAGGGAAGCGGC
CGCCCGGGTCGACTTCACAGGTGTTTGCGGCGTCTTTTGGAGTCTCCGGGCCTCAAGACG
CGGGGGCTGCTCTGCTCGCCCCACAGCCTTTCTTGTGCCCTCTGGTAGCCTCCCCATGCG
GAGAAATCGCCCCTCTGGTCCTCGCGGAAGTAGAGCTCCCTCCAGATGCCGCGATTCACC
TCTCCCAGCTCTTTAGCGGCTTGTTGCACGCCCCTAATTCTCCATTCCAGCCTTTCTTGG
AGGACCTCGGCTTGCAAAATCTGGCCCCTAATCCACCTATCCCTTCTGGAGGGTGTGTGC
TGGGTGGGACCGGGGCCGAGGTGTCTTCTGGCGATGCAGGTCTGGCTAGGAATCTTCTCC
TCGGGCAGGGACTGTCTCAGCACGCGGCACCACTGGTCCCCCTCCAGGGGGCCTTGTGGG
TCGATCTTCCACCAGTCGTTGCGGCGCTTCTCCTCTTTGCTCTCTTCCTTGAGGTTCATC

TCTTGATCCCTGGCCTCCTTGCTCTCAGCCATGGTGGCGAATTCTCGAGGCTAGCCTCCC
GGTGGTGGGTCGGTGGTCCCTGGGCAGGGGTCTCCAGATCCCGGACGAGCCCCCAAATGA
AAGACCCCCGAGACGGGTAGTCAATCACTCTGAGGAGACCCTCCCAAGGAACAGCGAGAC
CACGAGTCGGATGCAACAGCAAGAGGATTTATTGGATACACGGGTACCCGGGCGACTCAG
TCTATCGGAGGACTGGCGCGCCGAGTGAGGGGTTGTGAGCTCTTTTATAGAGCTCGGGAA
GCAGAAGCGCGCGAACAGAAGCGAGAAGCAGGCTGATTGGTTAATTCAAATAAGGCACAG
GGTCATTTCAGGTCCTTGGGGGAGCCTGGAAACATCTGATGGGTCTTAAGAAACTGCTGA
GGGTTGGGCCATATCTGGGGACCATCTGTTCTTGGCCCCGGGCCGGGGCCGAACCGCGGT
GACCATCTGTTCTTGGCCCCGGGCCGGGGCCGAAACTGCTCACCGCAGATATCCTGTTTG
GCCCAACGTTAGCTGTTTTCGTGTACCCGCCCTTGATCTGAACTTCTCTATTCTTGGTTT
GGTATTTTTCCATGCCTTGCAAAATGGCGTTACTGCGGCTATCAGGCTAAGCAATTTGAG
ATCTGGCCGAGGCGGCCTACTCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGT
TTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGG
CTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGG
GATAACGCAGGAAAGAACATGTATAACTTCGTATAATGTATGCTATACGAAGTTATACAT
GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTT
CCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCG
AAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTC
TCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGT
GGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAA
GCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA
TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAA
CTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT
TTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGAT
CTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCAT
GAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATC
AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC
ACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA
GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG
CAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGC
TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCAT
CGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG
GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT
CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA
TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAA
GTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGA
TAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGG
GCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGC
ACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGG
AAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACT
CTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACAT
ATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGT
GCCACCTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATC
AGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAG
ACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTG
GACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGATAAC

TTCGTATAATGTATGCTATACGAAGTTATCACTACGTGAACCATCACCCTAATCAAGTTT
TTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAG
AGCTTGACGGGGAAAGCCAACCTGGCTTATCGAAATTAATACGACTCACTATAGGGAGAC
CGGC

## pONY8.3G FB29 + (SEQ ID No 46)

AGATCTTGAATAATAAAATGTGTGTTTGTCCGAAATACGCGTTTTGAGATTTCTGTCGCC
GACTAAATTCATGTCGCGCGATAGTGGTGTTTATCGCCGATAGAGATGGCGATATTGGAA
AAATTGATATTTGAAAATATGGCATATTGAAAATGTCGCCGATGTGAGTTTCTGTGTAAC
TGATATCGCCATTTTTCCAAAAGTGATTTTTGGGCATACGCGATATCTGGCGATAGCGCT
TATATCGTTTACGGGGGATGGCGATAGACGACTTTGGTGACTTGGGCGATTCTGTGTGTC
GCAAATATCGCAGTTTCGATATAGGTGACAGACGATATGAGGCTATATCGCCGATAGAGG
CGACATCAAGCTGGCACATGGCCAATGCATATCGATCTATACATTGAATCAATATTGGCC
ATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCCATATCGTAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCC
ATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
TAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAAT
AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGG
ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT
GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTGCGATCGCCCGCC
CCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGT
TTAGTGAACCGGGCACTCAGATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGG
CCTTTGTAATAAATATAATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATC
CTACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGG
CTGATCGTAGGATCCCCGGGACAGCAGAGGAGAACTTACAGAAGTCTTCTGGAGGTGTTC
CTGGCCAGAACACAGGAGGACAGGTAAGATTGGGAGACCCTTTGACATTGGAGCAAGGCG
CTCAAGAAGTTAGAGAAGGTGACGGTACAAGGGTCTCAGAAATTAACTACTGGTAACTGT
AATTGGGCGCTAAGTCTAGTAGACTTATTTCATGATACCAACTTTGTAAAAGAAAAGGAC
TGGCAGCTGAGGGATGTCATTCCATTGCTGGAAGATGTAACTCAGACGCTGTCAGGACAA
GAAAGAGAGGCCTTTGAAAGAACATGGTGGGCAATTTCTGCTGTAAAGATGGGCCTCCAG
ATTAATAATGTAGTAGATGGAAAGGCATCATTCCAGCTCCTAAGAGCGAAATATGAAAAG
AAGACTGCTAATAAAAAGCAGTCTGAGCCCTCTGAAGAATATCTCTAGAACTAGTGGATC
CCCCGGGCTGCAGGAGTGGGGAGGCACGATGGCCGCTTTGGTCGAGGCGGATCCGGCCAT
TAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATA
CGTTGTATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCAT
GTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATA
GCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC
CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACG
TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGAT
AGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT
TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGC

AAATGGGCGGTAGGCATGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACC
GTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACC
GATCCAGCCTCCGCGGCCCCAAGCTTGTTGGGATCCACCGGTCGCCACCATGGTGAGCAA
GGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAA
CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGAC
CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCAC
CCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTT
CTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA
CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT
CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTA
CAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGT
GAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCA
GCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCAC
CCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCGA
CTCTAGAGTCGACCTGCAGGCATGCAAGCTTCAGCTGCTCGAGGGGGGGCCCGGTACCCA
GCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGGGAAGTATTTATCACTAATCAAGCAC
AAGTAATACATGAGAAACTTTTACTACAGCAAGCACAATCCTCCAAAAAATTTTGTTTTT
ACAAAATCCCTGGTGAACATGATTGGAAGGGACCTACTAGGGTGCTGTGGAAGGGTGATG
GTGCAGTAGTAGTTAATGATGAAGGAAAGGGAATAATTGCTGTACCATTAACCAGGACTA
AGTTACTAATAAAACCAAATTGAGTATTGTTGCAGGAAGCAAGACCCAACTACCATTGTC
AGCTGTGTTTCCTGACCTCAATATTTGTTATAAGGTTTGATATGAATCCCAGGGGGAATC
TCAACCCCTATTACCCAACAGTCAGAAAAATCTAAGTGTGAGGAGAACACAATGTTTCAA
CCTTATTGTTATAATAATGACAGTAAGAACAGCATGGCAGAATCGAAGGAAGCAAGAGAC
CAAGAATGAACCTGAAAGAAGAATCTAAAGAAGAAAAAAGAAGAAATGACTGGTGGAAAA
TAGGTATGTTTCTGTTATGCTTAGCAGGAACTACTGGAGGAATACTTTGGTGGTATGAAG
GACTCCCACAGCAACATTATATAGGGTTGGTGGCGATAGGGGGAAGATTAAACGGATCTG
GCCAATCAAATGCTATAGAATGCTGGGGTTCCTTCCCGGGGTGTAGACCATTTCAAAATT
ACTTCAGTTATGAGACCAATAGAAGCATGCATATGGATAATAATACTGCTACATTATTAG
AAGCTTTAACCAATATAACTGCTCTATAAATAACAAAACAGAATTAGAAACATGGAAGTT
AGTAAAGACTTCTGGCATAACTCCTTTACCTATTTCTTCTGAAGCTAACACTGGACTAAT
TAGACATAAGAGAGATTTTGGTATAAGTGCAATAGTGGCAGCTATTGTAGCCGCTACTGC
TATTGCTGCTAGCGCTACTATGTCTTATGTTGCTCTAACTGAGGTTAACAAAATAATGGA
AGTACAAAATCATACTTTTGAGGTAGAAAATAGTACTCTAAATGGTATGGATTTAATAGA
ACGACAAATAAAGATATTATATGCTATGATTCTTCAAACACATGCAGATGTTCAACTGTT
AAAGGAAAGACAACAGGTAGAGGAGACATTTAATTTAATTGGATGTATAGAAAGAACACA
TGTATTTTGTCATACTGGTCATCCCTGGAATATGTCATGGGGACATTTAAATGAGTCAAC
ACAATGGGATGACTGGGTAAGCAAAATGGAAGATTTAAATCAAGAGATACTAACTACACT
TCATGGAGCCAGGAACAATTTGGCACAATCCATGATAACATTCAATACACCAGATAGTAT
AGCTCAATTTGGAAAAGACCTTTGGAGTCATATTGGAAATTGGATTCCTGGATTGGGAGC
TTCCATTATAAAATATATAGTGATGTTTTTGCTTATTTATTTGTTACTAACCTCTTCGCC
TAAGATCCTCAGGGCCCTCTGGAAGGTGACCAGTGGTGCAGGGTCCTCCGGCAGTCGTTA
CCTGAAGAAAAAATTCCATCACAAACATGCATCGCGAGAAGACACCTGGGACCAGGCCCA
ACACAACATACACCTAGCAGGCGTGACCGGTGGATCAGGGGACAAATACTACAAGCAGAA
GTACTCCAGGAACGACTGGAATGGAGAATCAGAGGAGTACAACAGGCGGCCAAAGAGCTG
GGTGAAGTCAATCGAGGCATTTGGAGAGAGCTATATTTCCGAGAAGACCAAAGGGGAGAT
TTCTCAGCCTGGGGCGGCTATCAACGAGCACAAGAACGGCTCTGGGGGGAACAATCCTCA
CCAAGGGTCCTTAGACCTGGAGATTCGAAGCGAAGGAGGAAACATTTATGACTGTTGCAT
TAAAGCCCAAGAAGGAACTCTCGCTATCCCTTGCTGTGGATTTCCCTTATGGCTATTTTG

GGGACTAGTAATTATAGTAGGACGCATAGCAGGCTATGGATTACGTGGACTCGCTGTTAT
AATAAGGATTTGTATTAGAGGCTTAAATTTGATATTTGAAATAATCAGAAAAATGCTTGA
TTATATTGGAAGAGCTTTAAATCCTGGCACATCTCATGTATCAATGCCTCAGTATGTTTA
GAAAAACAAGGGGGGAACTGTGGGGTTTTTATGAGGGGTTTTATAAATGATTATAAGAGT
AAAAAGAAAGTTGCTGATGCTCTCATAACCTTGTATAACCCAAAGGACTAGCTCATGTTG
CTAGGCAACTAAACCGCAATAACCGCATTTGTGACGCGAGTTCCCCATTGGTGACGCGTT
AACTTCCTGTTTTTACAGTATATAAGTGCTTGTATTCTGACAATTGGGCACTCAGATTCT
GCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGGCCTTTGTAATAAATATAATTCTCTA
CTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATCCTACAGAGCTCATGCCTTGGCGTAA
TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA
CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA
ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGAGTAGGCCG
CCTCGGCCAGATCTCAAATTGCTTAGCCTGATAGCCGCAGTAACGCCATTTTGCAAGGCA
TGGAAAAATACCAAACCAAGAATAGAGAAGTTCAGATCAAGGGCGGGTACACGAAAACAG
CTAACGTTGGGCCAAACAGGATATCTGCGGTGAGCAGTTTCGGCCCCGGCCCGGGGCCAA
GAACAGATGGTCACCGCGGTTCGGCCCCGGCCCGGGGCCAAGAACAGATGGTCCCCAGAT
ATGGCCCAACCCTCAGCAGTTTCTTAAGACCCATCAGATGTTTCCAGGCTCCCCCAAGGA
CCTGAAATGACCCTGTGCCTTATTTGAATTAACCAATCAGCCTGCTTCTCGCTTCTGTTC
GCGCGCTTCTGCTTCCCGAGCTCTATAAAAGAGCTCACAACCCCTCACTCGGCGCGCCAG
TCCTCCGATAGACTGAGTCGCCCGGGTACCCGTGTATCCAATAAATCCTCTTGCTGTTGC
ATCCGACTCGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCAGAGTGATTGACTACCCGT
CTCGGGGGTCTTTCATTTGGGGGCTCGTCCGGGATCTGGAGACCCCTGCCCAGGGACCAC
CGACCCACCACCGGGAGGCTAGCCTCGAGAATTCGCCACCATGGCTGAGAGCAAGGAGGC
CAGGGATCAAGAGATGAACCTCAAGGAAGAGAGCAAAGAGGAGAAGCGCCGCAACGACTG
GTGGAAGATCGACCCACAAGGCCCCCTGGAGGGGGACCAGTGGTGCCGCGTGCTGAGACA
GTCCCTGCCCGAGGAGAAGATTCCTAGCCAGACCTGCATCGCCAGAAGACACCTCGGCCC
CGGTCCCACCCAGCACACACCCTCCAGAAGGGATAGGTGGATTAGGGGCCAGATTTTGCA
AGCCGAGGTCCTCCAAGAAAGGCTGGAATGGAGAATTAGGGGCGTGCAACAAGCCGCTAA
AGAGCTGGGAGAGGTGAATCGCGGCATCTGGAGGGAGCTCTACTTCCGCGAGGACCAGAG
GGGCGATTTCTCCGCATGGGGAGGCTACCAGAGGGCACAAGAAAGGCTGTGGGGCGAGCA
GAGCAGCCCCCGCGTCTTGAGGCCCGGAGACTCCAAAAGACGCCGCAAACACCTGTGAAG
TCGACCCGGGCGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGA
TACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGT
GAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAC
AACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAA
AGCAAGTAAAACCTCTACAAATGTGGTAAAATCCGATAAGGATCGATCCGGGCTGGCGTA
ATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAAT
GGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGAC
CGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGC
CACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATT
TAGAGCTTTACGGCACCTCGACCGCAAAAAACTTGATTTGGGTGATGGTTCACGTAGGCC
GCCTCGGCCGCCCGGGCATCACTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGT
TTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGG
CTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGG
GATAACGCAGGAAAGAACATGTATAACTTCGTATAATGTATGCTATACGAAGTTATACAT
GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTT
CCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCG
AAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTC
TCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGT

GGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAA
GCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA
TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAA
CTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT
TTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGAT
CTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCAT
GAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATC
AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC
ACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA
GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG
CAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGC
TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCAT
CGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG
GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT
CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA
TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAA
GTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGA
TAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGG
GCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGC
ACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGG
AAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACT
CTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACAT
ATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGT
GCCACCTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATC
AGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAG
ACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTG
GACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGATAAC
TTCGTATAATGTATGCTATACGAAGTTATCACTACGTGAACCATCACCCTAATCAAGTTT
TTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAG
AGCTTGACGGGGAAAGCCAACCTGGCTTATCGAAATTAATACGACTCACTATAGGGAGAC
CGGC

## pONY8.3GPGK – (SEQ ID No 47)

AGATCTTGAATAATAAAATGTGTGTTTGTCCGAAATACGCGTTTTGAGATTTCTGTCGCC
GACTAAATTCATGTCGCGCGATAGTGGTGTTTATCGCCGATAGAGATGGCGATATTGGAA
AAATTGATATTTGAAAATATGGCATATTGAAAATGTCGCCGATGTGAGTTTCTGTGTAAC
TGATATCGCCATTTTTCCAAAAGTGATTTTTGGGCATACGCGATATCTGGCGATAGCGCT
TATATCGTTTACGGGGGATGGCGATAGACGACTTTGGTGACTTGGGCGATTCTGTGTGTC
GCAAATATCGCAGTTTCGATATAGGTGACAGACGATATGAGGCTATATCGCCGATAGAGG
CGACATCAAGCTGGCACATGGCCAATGCATATCGATCTATACATTGAATCAATATTGGCC
ATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCCATATCGTAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCC
ATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
TAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAAT

AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGG
ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT
GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTGCGATCGCCCGCC
CCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGT
TTAGTGAACCGGGCACTCAGATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGG
CCTTTGTAATAAATATAATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATC
CTACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGG
CTGATCGTAGGATCCCCGGGACAGCAGAGGAGAACTTACAGAAGTCTTCTGGAGGTGTTC
CTGGCCAGAACACAGGAGGACAGGTAAGATTGGGAGACCCTTTGACATTGGAGCAAGGCG
CTCAAGAAGTTAGAGAAGGTGACGGTACAAGGGTCTCAGAAATTAACTACTGGTAACTGT
AATTGGGCGCTAAGTCTAGTAGACTTATTTCATGATACCAACTTTGTAAAAGAAAAGGAC
TGGCAGCTGAGGGATGTCATTCCATTGCTGGAAGATGTAACTCAGACGCTGTCAGGACAA
GAAAGAGAGGCCTTTGAAAGAACATGGTGGGCAATTTCTGCTGTAAAGATGGGCCTCCAG
ATTAATAATGTAGTAGATGGAAAGGCATCATTCCAGCTCCTAAGAGCGAAATATGAAAAG
AAGACTGCTAATAAAAAGCAGTCTGAGCCCTCTGAAGAATATCTCTAGAACTAGTGGATC
CCCCGGGCTGCAGGAGTGGGGAGGCACGATGGCCGCTTTGGTCGAGGCGGATCCGGCCAT
TAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATA
CGTTGTATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCAT
GTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATA
GCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC
CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACG
TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGAT
AGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT
TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGC
AAATGGGCGGTAGGCATGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACC
GTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACC
GATCCAGCCTCCGCGGCCCCAAGCTTGTTGGGATCCACCGGTCGCCACCATGGTGAGCAA
GGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAA
CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGAC
CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCAC
CCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTT
CTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA
CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT
CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTA
CAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGT
GAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCA
GCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCAC
CCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCGA
CTCTAGAGTCGACCTGCAGGCATGCAAGCTTCAGCTGCTCGAGGGGGGGCCCGGTACCCA
GCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGGGAAGTATTTATCACTAATCAAGCAC
AAGTAATACATGAGAAACTTTTACTACAGCAAGCACAATCCTCCAAAAAATTTTGTTTTT
ACAAAATCCCTGGTGAACATGATTGGAAGGGACCTACTAGGGTGCTGTGGAAGGGTGATG

GTGCAGTAGTAGTTAATGATGAAGGAAAGGGAATAATTGCTGTACCATTAACCAGGACTA
AGTTACTAATAAAACCAAATTGAGTATTGTTGCAGGAAGCAAGACCCAACTACCATTGTC
AGCTGTGTTTCCTGACCTCAATATTTGTTATAAGGTTTGATATGAATCCCAGGGGGAATC
TCAACCCCTATTACCCAACAGTCAGAAAAAATCTAAGTGTGAGGAGAACACAATGTTTCAA
CCTTATTGTTATAATAATGACAGTAAGAACAGCATGGCAGAATCGAAGGAAGCAAGAGAC
CAAGAATGAACCTGAAAGAAGAATCTAAAGAAGAAAAAAGAAGAAATGACTGGTGGAAAA
TAGGTATGTTTCTGTTATGCTTAGCAGGAACTACTGGAGGAATACTTTGGTGGTATGAAG
GACTCCCACAGCAACATTATATAGGGTTGGTGGCGATAGGGGGAAGATTAAACGGATCTG
GCCAATCAAATGCTATAGAATGCTGGGGTTCCTTCCCGGGGTGTAGACCATTTCAAAATT
ACTTCAGTTATGAGACCAATAGAAGCATGCATATGGATAATAATACTGCTACATTATTAG
AAGCTTTAACCAATATAACTGCTCTATAAATAACAAAACAGAATTAGAAACATGGAAGTT
AGTAAAGACTTCTGGCATAACTCCTTTACCTATTTCTTCTGAAGCTAACACTGGACTAAT
TAGACATAAGAGAGATTTTGGTATAAGTGCAATAGTGGCAGCTATTGTAGCCGCTACTGC
TATTGCTGCTAGCGCTACTATGTCTTATGTTGCTCTAACTGAGGTTAACAAAATAATGGA
AGTACAAAATCATACTTTTGAGGTAGAAAATAGTACTCTAAATGGTATGGATTTAATAGA
ACGACAAATAAAGATATTATATGCTATGATTCTTCAAACACATGCAGATGTTCAACTGTT
AAAGGAAAGACAACAGGTAGAGGAGACATTTAATTTAATTGGATGTATAGAAAGAACACA
TGTATTTTGTCATACTGGTCATCCCTGGAATATGTCATGGGGACATTTAAATGAGTCAAC
ACAATGGGATGACTGGGTAAGCAAAATGGAAGATTTAAATCAAGAGATACTAACTACACT
TCATGGAGCCAGGAACAATTTGGCACAATCCATGATAACATTCAATACACCAGATAGTAT
AGCTCAATTTGGAAAAGACCTTTGGAGTCATATTGGAAATTGGATTCCTGGATTGGGAGC
TTCCATTATAAAATATATAGTGATGTTTTTGCTTATTTATTTGTTACTAACCTCTTCGCC
TAAGATCCTCAGGGCCCTCTGGAAGGTGACCAGTGGTGCAGGGTCCTCCGGCAGTCGTTA
CCTGAAGAAAAAATTCCATCACAAACATGCATCGCGAGAAGACACCTGGGACCAGGCCCA
ACACAACATACACCTAGCAGGCGTGACCGGTGGATCAGGGGACAAATACTACAAGCAGAA
GTACTCCAGGAACGACTGGAATGGAGAATCAGAGGAGTACAACAGGCGGCCAAAGAGCTG
GGTGAAGTCAATCGAGGCATTTGGAGAGAGCTATATTTCCGAGAAGACCAAAGGGGAGAT
TTCTCAGCCTGGGGCGGCTATCAACGAGCACAAGAACGGCTCTGGGGGGAACAATCCTCA
CCAAGGGTCCTTAGACCTGGAGATTCGAAGCGAAGGAGGAAACATTTATGACTGTTGCAT
TAAAGCCCAAGAAGGAACTCTCGCTATCCCTTGCTGTGGATTTCCCTTATGGCTATTTTG
GGGACTAGTAATTATAGTAGGACGCATAGCAGGCTATGGATTACGTGGACTCGCTGTTAT
AATAAGGATTTGTATTAGAGGCTTAAATTTGATATTTGAAATAATCAGAAAAATGCTTGA
TTATATTGGAAGAGCTTTAAATCCTGGCACATCTCATGTATCAATGCCTCAGTATGTTTA
GAAAAACAAGGGGGGAACTGTGGGGTTTTTATGAGGGGTTTTATAAATGATTATAAGAGT
AAAAAGAAAGTTGCTGATGCTCTCATAACCTTGTATAACCCAAAGGACTAGCTCATGTTG
CTAGGCAACTAAACCGCAATAACCGCATTTGTGACGCGAGTTCCCCATTGGTGACGCGTT
AACTTCCTGTTTTTACAGTATATAAGTGCTTGTATTCTGACAATTGGGCACTCAGATTCT
GCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGGCCTTTGTAATAAATATAATTCTCTA
CTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATCCTACAGAGCTCATGCCTTGGCGTAA
TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA
CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA
ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGTGATGCCCG
GGCGGCCGAGGCGGCCTACGTGAACCATCACCCAAATCAAGTTTTTTGCGGTCGAGGTGC
CGTAAAGCTCTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAG
CCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTG
GCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTA
CAGGGCGCGTCCATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGG
CCTCTTCGCTATTACGCCAGCCCGGATCGATCCTTATCGGATTTTACCACATTTGTAGAG
GTTTTACTTGCTTTAAAAAACCTCCCACATCTCCCCCTGAACCTGAAACATAAAATGAAT

GCAATTGTTGTTGTTAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGC
ATCACAAATTTCACAAATAAAGCATTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAAA
CTCATCAATGTATCTTATCATGTCTGCTCGAAGCATTAACCCTCACTAAAGGGAAGCGGC
CGCCCGGGTCGACTTCACAGGTGTTTGCGGCGTCTTTTGGAGTCTCCGGGCCTCAAGACG
CGGGGGCTGCTCTGCTCGCCCCACAGCCTTTCTTGTGCCCTCTGGTAGCCTCCCCATGCG
GAGAAATCGCCCCTCTGGTCCTCGCGGAAGTAGAGCTCCCTCCAGATGCCGCGATTCACC
TCTCCCAGCTCTTTAGCGGCTTGTTGCACGCCCCTAATTCTCCATTCCAGCCTTTCTTGG
AGGACCTCGGCTTGCAAAATCTGGCCCCTAATCCACCTATCCCTTCTGGAGGGTGTGTGC
TGGGTGGGACCGGGGCCGAGGTGTCTTCTGGCGATGCAGGTCTGGCTAGGAATCTTCTCC
TCGGGCAGGGACTGTCTCAGCACGCGGCACCACTGGTCCCCCTCCAGGGGGCCTTGTGGG
TCGATCTTCCACCAGTCGTTGCGGCGCTTCTCCTCTTTGCTCTCTTCCTTGAGGTTCATC
TCTTGATCCCTGGCCTCCTTGCTCTCAGCCATGGTGGCGAATTCTCGAGGCTAGCCTGGG
GAGAGAGGTCGGTGATTCGGTCAACGAGGGAGCCGACTGCCGACGTGCGCTCCGGAGGCT
TGCAGAATGCGGAACACCGCGCGGGCAGGAACAGGGCCCACACTACCGCCCCACACCCCG
CCTCCCGCACCGCCCCTTCCCGGCCGCTGCTCTCGGCGCGCCCCGCTGAGCAGCCGCTAT
TGGCCACAGCCCATCGCGGTCGGCGCGCTGCCATTGCTCCCTGGCGCTGTCCGTCTGCGA
GGGTACTAGTGAGACGTGCGGCTTCCGTTTGTCACGTCCGGCACGCCGCGAACCGCAAGG
AACCTTCCCGACTTAGGGGCGGAGCAGGAAGCGTCGCCGGGGGGCCCACAAGGGTAGCGG
CGAAGATCCGGGTGACGCTGCGAACGGACGTGAAGAATGTGCGAGACCCAGGGTCGGCGC
CGCTGCGTTTCCCGGAACCACGCCCAGAGCAGCCGCGTCCCTGCGCAAACCCAGGGCTGC
CTTGGAAAAGGCGCAACCCCAACCCCAGATCTGGCCGAGGCGGCCTACTCTGCATTAATG
AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT
CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGC
GGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTATAACTTCGT
ATAATGTATGCTATACGAAGTTATACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC
GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACA
AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATC
TCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGC
CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACT
TATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTG
CTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTA
TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCA
AACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAA
AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACG
AAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCC
TTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTG
ACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCAT
CCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAA
TAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCA
TCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGC
GCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTT
CATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAA
AAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTAT
CACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCT
TTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGA
GTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAG

TGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGA
GATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCA
CCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGG
CGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATC
AGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAG
GGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTAAATTGTAAGCGTTAATATTTTGTT
AAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGG
CAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTG
GAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTA
TCAGGGCGATGGCCCACTACGTGATAACTTCGTATAATGTATGCTATACGAAGTTATCAC
TACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATC
GGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCAACCTGGCTTATCG
AAATTAATACGACTCACTATAGGGAGACCGGC

## pONY8.3G PGK + (SEQ ID No 48)

AGATCTTGAATAATAAAATGTGTGTTTGTCCGAAATACGCGTTTTGAGATTTCTGTCGCC
GACTAAATTCATGTCGCGCGATAGTGGTGTTTATCGCCGATAGAGATGGCGATATTGGAA
AAATTGATATTTGAAAATATGGCATATTGAAAATGTCGCCGATGTGAGTTTCTGTGTAAC
TGATATCGCCATTTTTCCAAAAGTGATTTTTGGGCATACGCGATATCTGGCGATAGCGCT
TATATCGTTTACGGGGGATGGCGATAGACGACTTTGGTGACTTGGGCGATTCTGTGTGTC
GCAAATATCGCAGTTTCGATATAGGTGACAGACGATATGAGGCTATATCGCCGATAGAGG
CGACATCAAGCTGGCACATGGCCAATGCATATCGATCTATACATTGAATCAATATTGGCC
ATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCCATATCGTAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCC
ATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
TAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAAT
AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGG
ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT
GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTGCGATCGCCCGCC
CCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGT
TTAGTGAACCGGGCACTCAGATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGG
CCTTTGTAATAAATATAATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATC
CTACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGG
CTGATCGTAGGATCCCCGGGACAGCAGAGGAGAACTTACAGAAGTCTTCTGGAGGTGTTC
CTGGCCAGAACACAGGAGGACAGGTAAGATTGGGAGACCCTTTGACATTGGAGCAAGGCG
CTCAAGAAGTTAGAGAAGGTGACGGTACAAGGGTCTCAGAAATTAACTACTGGTAACTGT
AATTGGGCGCTAAGTCTAGTAGACTTATTTCATGATACCAACTTTGTAAAAGAAAAGGAC
TGGCAGCTGAGGGATGTCATTCCATTGCTGGAAGATGTAACTCAGACGCTGTCAGGACAA
GAAAGAGAGGCCTTTGAAAGAACATGGTGGGCAATTTCTGCTGTAAAGATGGGCCTCCAG
ATTAATAATGTAGTAGATGGAAAGGCATCATTCCAGCTCCTAAGAGCGAAATATGAAAAG
AAGACTGCTAATAAAAAGCAGTCTGAGCCCTCTGAAGAATATCTCTAGAACTAGTGGATC
CCCCGGGCTGCAGGAGTGGGGAGGCACGATGGCCGCTTTGGTCGAGGCGGATCCGGCCAT
TAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATA
CGTTGTATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCAT
GTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATA
GCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGC
CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG
GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG
CCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACG
TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGAT
AGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT
TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGC
AAATGGGCGGTAGGCATGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACC
GTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACC
GATCCAGCCTCCGCGGCCCCAAGCTTGTTGGGATCCACCGGTCGCCACCATGGTGAGCAA
GGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAA
CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGAC

CCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCAC
CCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTT
CTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGA
CGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCAT
CGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTA
CAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGT
GAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCA
GCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCAC
CCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTT
CGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCGA
CTCTAGAGTCGACCTGCAGGCATGCAAGCTTCAGCTGCTCGAGGGGGGGCCCGGTACCCA
GCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGGGAAGTATTTATCACTAATCAAGCAC
AAGTAATACATGAGAAACTTTTACTACAGCAAGCACAATCCTCCAAAAAATTTTGTTTTT
ACAAAATCCCTGGTGAACATGATTGGAAGGGACCTACTAGGGTGCTGTGGAAGGGTGATG
GTGCAGTAGTAGTTAATGATGAAGGAAAGGGAATAATTGCTGTACCATTAACCAGGACTA
AGTTACTAATAAAACCAAATTGAGTATTGTTGCAGGAAGCAAGACCCAACTACCATTGTC
AGCTGTGTTTCCTGACCTCAATATTTGTTATAAGGTTTGATATGAATCCCAGGGGGAATC
TCAACCCCTATTACCCAACAGTCAGAAAAATCTAAGTGTGAGGAGAACACAATGTTTCAA
CCTTATTGTTATAATAATGACAGTAAGAACAGCATGGCAGAATCGAAGGAAGCAAGAGAC
CAAGAATGAACCTGAAAGAAGAATCTAAAGAAGAAAAAAGAAGAAATGACTGGTGGAAAA
TAGGTATGTTTCTGTTATGCTTAGCAGGAACTACTGGAGGAATACTTTGGTGGTATGAAG
GACTCCCACAGCAACATTATATAGGGTTGGTGGCGATAGGGGGAAGATTAAACGGATCTG
GCCAATCAAATGCTATAGAATGCTGGGGTTCCTTCCCGGGGTGTAGACCATTTCAAAATT
ACTTCAGTTATGAGACCAATAGAAGCATGCATATGGATAATAATACTGCTACATTATTAG
AAGCTTTAACCAATATAACTGCTCTATAAATAACAAAACAGAATTAGAAACATGGAAGTT
AGTAAAGACTTCTGGCATAACTCCTTTACCTATTTCTTCTGAAGCTAACACTGGACTAAT
TAGACATAAGAGAGATTTTGGTATAAGTGCAATAGTGGCAGCTATTGTAGCCGCTACTGC
TATTGCTGCTAGCGCTACTATGTCTTATGTTGCTCTAACTGAGGTTAACAAAATAATGGA
AGTACAAAATCATACTTTTGAGGTAGAAAATAGTACTCTAAATGGTATGGATTTAATAGA
ACGACAAATAAAGATATTATATGCTATGATTCTTCAAACACATGCAGATGTTCAACTGTT
AAAGGAAAGACAACAGGTAGAGGAGACATTTAATTTAATTGGATGTATAGAAAGAACACA
TGTATTTTGTCATACTGGTCATCCCTGGAATATGTCATGGGGACATTTAAATGAGTCAAC
ACAATGGGATGACTGGGTAAGCAAAATGGAAGATTTAAATCAAGAGATACTAACTACACT
TCATGGAGCCAGGAACAATTTGGCACAATCCATGATAACATTCAATACACCAGATAGTAT
AGCTCAATTTGGAAAAGACCTTTGGAGTCATATTGGAAATTGGATTCCTGGATTGGGAGC
TTCCATTATAAAATATATAGTGATGTTTTTGCTTATTTATTTGTTACTAACCTCTTCGCC
TAAGATCCTCAGGGCCCTCTGGAAGGTGACCAGTGGTGCAGGGTCCTCCGGCAGTCGTTA
CCTGAAGAAAAAATTCCATCACAAACATGCATCGCGAGAAGACACCTGGGACCAGGCCCA
ACACAACATACACCTAGCAGGCGTGACCGGTGGATCAGGGGACAAATACTACAAGCAGAA
GTACTCCAGGAACGACTGGAATGGAGAATCAGAGGAGTACAACAGGCGGCCAAAGAGCTG
GGTGAAGTCAATCGAGGCATTTGGAGAGAGCTATATTTCCGAGAAGACCAAAGGGGAGAT
TTCTCAGCCTGGGGCGGCTATCAACGAGCACAAGAACGGCTCTGGGGGGAACAATCCTCA
CCAAGGGTCCTTAGACCTGGAGATTCGAAGCGAAGGAGGAAACATTTATGACTGTTGCAT
TAAAGCCCAAGAAGGAACTCTCGCTATCCCTTGCTGTGGATTTCCCTTATGGCTATTTTG
GGGACTAGTAATTATAGTAGGACGCATAGCAGGCTATGGATTACGTGGACTCGCTGTTAT
AATAAGGATTTGTATTAGAGGCTTAAATTTGATATTTGAAATAATCAGAAAAATGCTTGA
TTATATTGGAAGAGCTTTAAATCCTGGCACATCTCATGTATCAATGCCTCAGTATGTTTA
GAAAAACAAGGGGGGAACTGTGGGGTTTTTATGAGGGGTTTTATAAATGATTATAAGAGT
AAAAAGAAAGTTGCTGATGCTCTCATAACCTTGTATAACCCAAAGGACTAGCTCATGTTG

CTAGGCAACTAAACCGCAATAACCGCATTTGTGACGCGAGTTCCCCATTGGTGACGCGTT
AACTTCCTGTTTTTACAGTATATAAGTGCTTGTATTCTGACAATTGGGCACTCAGATTCT
GCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGGCCTTTGTAATAAATATAATTCTCTA
CTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATCCTACAGAGCTCATGCCTTGGCGTAA
TCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCCACACAACATA
CGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTCACATTA
ATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGAGTAGGCCG
CCTCGGCCAGATCTGGGGTTGGGGTTGCGCCTTTTCCAAGGCAGCCCTGGGTTTGCGCAG
GGACGCGGCTGCTCTGGGCGTGGTTCCGGGAAACGCAGCGGCGCCGACCCTGGGTCTCGC
ACATTCTTCACGTCCGTTCGCAGCGTCACCCGGATCTTCGCCGCTACCCTTGTGGGCCCC
CCGGCGACGCTTCCTGCTCCGCCCCTAAGTCGGGAAGGTTCCTTGCGGTTCGCGGCGTGC
CGGACGTGACAAACGGAAGCCGCACGTCTCACTAGTACCCTCGCAGACGGACAGCGCCAG
GGAGCAATGGCAGCGCGCCGACCGCGATGGGCTGTGGCCAATAGCGGCTGCTCAGCGGGG
CGCGCCGAGAGCAGCGGCCGGGAAGGGGCGGTGCGGGAGGCGGGGTGTGGGGCGGTAGTG
TGGGCCCTGTTCCTGCCCGCGCGGTGTTCCGCATTCTGCAAGCCTCCGGAGCGCACGTCG
GCAGTCGGCTCCCTCGTTGACCGAATCACCGACCTCTCTCCCCAGGCTAGCCTCGAGAAT
TCGCCACCATGGCTGAGAGCAAGGAGGCCAGGGATCAAGAGATGAACCTCAAGGAAGAGA
GCAAAGAGGAGAAGCGCCGCAACGACTGGTGGAAGATCGACCCACAAGGCCCCCTGGAGG
GGGACCAGTGGTGCCGCGTGCTGAGACAGTCCCTGCCCGAGGAGAAGATTCCTAGCCAGA
CCTGCATCGCCAGAAGACACCTCGGCCCCGGTCCCACCCAGCACACACCCTCCAGAAGGG
ATAGGTGGATTAGGGGCCAGATTTTGCAAGCCGAGGTCCTCCAAGAAAGGCTGGAATGGA
GAATTAGGGGCGTGCAACAAGCCGCTAAAGAGCTGGGAGAGGTGAATCGCGGCATCTGGA
GGGAGCTCTACTTCCGCGAGGACCAGAGGGGCGATTTCTCCGCATGGGGAGGCTACCAGA
GGGCACAAGAAAGGCTGTGGGGCGAGCAGAGCAGCCCCCGCGTCTTGAGGCCCGGAGACT
CCAAAAGACGCCGCAAACACCTGTGAAGTCGACCCGGGCGGCCGCTTCCCTTTAGTGAGG
GTTAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACTAG
AATGCAGTGAAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAC
CATTATAAGCTGCAATAAACAAGTTAACAACAACAATTGCATTCATTTTATGTTTCAGGT
TCAGGGGGAGATGTGGGAGGTTTTTTAAAGCAAGTAAAACCTCTACAAATGTGGTAAAAT
CCGATAAGGATCGATCCGGGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCC
CAACAGTTGCGCAGCCTGAATGGCGAATGGACGCGCCCTGTAGCGGCGCATTAAGCGCGG
CGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTC
CTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAA
ATCGGGGGCTCCCTTTAGGGTTCCGATTTAGAGCTTTACGGCACCTCGACCGCAAAAAAC
TTGATTTGGGTGATGGTTCACGTAGGCCGCCTCGGCCGCCCGGGCATCACTGCATTAATG
AATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCT
CACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGC
GGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTATAACTTCGT
ATAATGTATGCTATACGAAGTTATACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC
GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACA
AAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGT
TTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACC
TGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATC
TCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGC
CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACT
TATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTG
CTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTA
TCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCA
AACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAA

AAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACG
AAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCC
TTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTG
ACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCAT
CCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTG
GCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAA
TAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCA
TCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGC
GCAACGTTGTTGCCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTT
CATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAA
AAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTAT
CACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCT
TTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGA
GTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAG
TGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGA
GATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCA
CCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGG
CGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATC
AGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAG
GGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTAAATTGTAAGCGTTAATATTTTGTT
AAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGG
CAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTG
GAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTA
TCAGGGCGATGGCCCACTACGTGATAACTTCGTATAATGTATGCTATACGAAGTTATCAC
TACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATC
GGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCAACCTGGCTTATCG
AAATTAATACGACTCACTATAGGGAGACCGGC

SEQ ID No 51

## pONY3.2IRESHyg

AGATCTCCCGATCCCCTATGGTCGACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTA
AGCCAGTATCTGCTCCCTGCTTGTGTGTTGGAGGTCGCTGAGTAGTGCGCGAGCAAAATT
TAAGCTACAACAAGGCAAGGCTTGACCGACAATTGCATGAAGAATCTGCTTAGGGTTAGG
CGTTTTGCGCTGCTTCGCGATGTACGGGCCAGATATACGCGTTGACATTGATTATTGACT
AGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGC
GTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTG
ACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAA
TGGGTGGACTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCA
AGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC
ATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACC
ATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA
TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG
GACTTTCCAAAATGTCGTAACAACTGCGATCGCCCGCCCCGTTGACGCAAATGGGCGGTA
GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCACTA
GAAGCTTTATTGCGGTAGTTTATCACAGTTAAATTGCTAACGCAGTCAGTGCTTCTGACA
CAACAGTCTCGAACTTAAGCTGCAGTGACTCTCTTAAGGTAGCCTTGCAGAAGTTGGTCG
TGAGGCACTGGGCAGGTAAGTATCAAGGTTACAAGACAGGTTTAAGGAGACCAATAGAAA
CTGGGCTTGTCGAGACAGAGAAGACTCTTGCGTTTCTGATAGGCACCTATTGGTCTTACT
GACATCCACTTTGCCTTTCTCTCCACAGGTGTCCACTCCCAGTTCAATTACAGCTCTTAA
GGCTAGAGTACTTAATACGACTCACTATAGGCTAGCCTCGAGGTCGACGGTATCGCCCGA
ACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGGCTGATCGTAGGATCCCC
GGGACAGCAGAGGAGAACTTACAGAAGTCTTCTGGAGGTGTTCCTGGCCAGAACACAGGA
GGACAGGTAAGATGGGAGACCCTTTGACATGGAGCAAGGCGCTCAAGAAGTTAGAGAAGG
TGACGGTACAAGGGTCTCAGAAATTAACTACTGGTAACTGTAATTGGGCGCTAAGTCTAG
TAGACTTATTTCATGATACCAACTTTGTAAAAGAAAAGGACTGGCAGCTGAGGGATGTCA
TTCCATTGCTGGAAGATGTAACTCAGACGCTGTCAGGACAAGAAAGAGAGGCCTTTGAAA
GAACATGGTGGGCAATTTCTGCTGTAAAGATGGGCCTCCAGATTAATAATGTAGTAGATG
GAAAGGCATCATTCCAGCTCCTAAGAGCGAAATATGAAAAGAAGACTGCTAATAAAAAGC
AGTCTGAGCCCTCTGAAGAATATCCAATCATGATAGATGGGGCTGGAAACAGAAATTTTA
GACCTCTAACACCTAGAGGATATACTACTTGGGTGAATACCATACAGACAAATGGTCTAT
TAAATGAAGCTAGTCAAAACTTATTTGGGATATTATCAGTAGACTGTACTTCTGAAGAAA
TGAATGCATTTTTGGATGTGGTACCTGGCCAGGCAGGACAAAAGCAGATATTACTTGATG
CAATTGATAAGATAGCAGATGATTGGGATAATAGACATCCATTACCGAATGCTCCACTGG
TGGCACCACCACAAGGGCCTATTCCCATGACAGCAAGGTTTATTAGAGGTTTAGGAGTAC
CTAGAGAAAGACAGATGGAGCCTGCTTTTGATCAGTTTAGGCAGACATATAGACAATGGA
TAATAGAAGCCATGTCAGAAGGCATCAAAGTGATGATTGGAAAACCTAAAGCTCAAAATA
TTAGGCAAGGAGCTAAGGAACCTTACCCAGAATTTGTAGACAGACTATTATCCCAAATAA
AAAGTGAGGGACATCCACAAGAGATTTCAAAATTCTTGACTGATACACTGACTATTCAGA
ACGCAAATGAGGAATGTAGAAATGCTATGAGACATTTAAGACCAGAGGATACATTAGAAG
AGAAAATGTATGCTTGCAGAGACATTGGAACTACAAAACAAAAGATGATGTTATTGGCAA
AAGCACTTCAGACTGGTCTTGCGGGCCCATTTAAAGGTGGAGCCTTGAAAGGAGGGCCAC
TAAAGGCAGCACAAACATGTTATAACTGTGGGAAGCCAGGACATTTATCTAGTCAATGTA
GAGCACCTAAAGTCTGTTTTAAATGTAAACAGCCTGGACATTTCTCAAAGCAATGCAGAA
GTGTTCCAAAAAACGGGAAGCAAGGGGCTCAAGGGAGGCCCCAGAAACAAACTTTCCCGA
TACAACAGAAGAGTCAGCACAACAAATCTGTTGTACAAGAGACTCCTCAGACTCAAAATC
TGTACCCAGATCTGAGCGAAATAAAAAAGGAATACAATGTCAAGGAGAAGGATCAAGTAG
AGGATCTCAACCTGGACAGTTTGTGGGAGTAACATATAATCTAGAGAAAAGGCCTACTAC
AATAGTATTAATTAATGATACTCCCTTAAATGTACTGTTAGACACAGGAGCAGATACTTC
AGTGTTGACTACTGCACATTATAATAGGTTAAAATATAGAGGGAGAAAATATCAAGGGAC
GGGAATAATAGGAGTGGGAGGAAATGTGGAAACATTTTCTACGCCTGTGACTATAAAGAA
AAAGGGTAGACACATTAAGACAAGAATGCTAGTGGCAGATATTCCAGTGACTATTTTGGG
ACGAGATATTCTTCAGGACTTAGGTGCAAAATTGGTTTTGGCACAGCTCTCCAAGGAAAT
AAAATTTAGAAAAATAGAGTTAAAAGAGGGCACAATGGGGCCAAAAATTCCTCAATGGCC
ACTCACTAAGGAGAAACTAGAAGGGGCCAAAGAGATAGTCCAAAGACTATTGTCAGAGGG
AAAAATATCAGAAGCTAGTGACAATAATCCTTATAATTCACCCATATTTGTAATAAAAAA
GAGGTCTGGCAAATGGAGGTTATTACAAGATCTGAGAGAATTAAACAAAACAGTACAAGT
AGGAACGGAAATATCCAGAGGATTGCCTCACCCGGGAGGATTAATTAAATGTAAACACAT
GACTGTATTAGATATTGGAGATGCATATTTCACTATACCCTTAGATCCAGAGTTTAGACC
ATATACAGCTTTCACTATTCCCTCCATTAATCATCAAGAACCAGATAAAAGATATGTGTG
GAAATGTTTACCACAAGGATTCGTGTTGAGCCCATATATATATCAGAAAACATTACAGGA
AATTTTACAACCTTTTAGGGAAAGATATCCTGAAGTACAATTGTATCAATATATGGATGA

TTTGTTCATGGGAAGTAATGGTTCTAAAAAACAACACAAAGAGTTAATCATAGAATTAAG
GGCGATCTTACTGGAAAAGGGTTTTGAGACACCAGATGATAAATTACAAGAAGTGCCACC
TTATAGCTGGCTAGGTTATCAACTTTGTCCTGAAAATTGGAAAGTACAAAAAATGCAATT
AGACATGGTAAAGAATCCAACCCTTAATGATGTGCAAAAATTAATGGGGAATATAACATG
GATGAGCTCAGGGATCCCAGGGTTGACAGTAAAACACATTGCAGCTACTACTAAGGGATG
TTTAGAGTTGAATCAAAAAGTAATTTGGACGGAAGAGGCACAAAAAGAGTTAGAAGAAAA
TAATGAGAAGATTAAAAATGCTCAAGGGTTACAATATTATAATCCAGAAGAAGAAATGTT
ATGTGAGGTTGAAATTACAAAAAATTATGAGGCAACTTATGTTATAAAACAATCACAAGG
AATCCTATGGGCAGGTAAAAAGATTATGAAGGCTAATAAGGGATGGTCAACAGTAAAAAA
TTTAATGTTATTGTTGCAACATGTGGCAACAGAAAGTATTACTAGAGTAGGAAAATGTCC
AACGTTTAAGGTACCATTTACCAAAGAGCAAGTAATGTGGGAAATGCAAAAAGGATGGTA
TTATTCTTGGCTCCCAGAAATAGTATATACACATCAAGTAGTTCATGATGATTGGAGAAT
GAAATTGGTAGAAGAACCTACATCAGGAATAACAATATACACTGATGGGGGAAAACAAAA
TGGAGAAGGAATAGCAGCTTATGTGACCAGTAATGGGAGAACTAAACAGAAAAGGTTAGG
ACCTGTCACTCATCAAGTTGCTGAAAGAATGGCAATACAAATGGCATTAGAGGATACCAG
AGATAAACAAGTAAATATAGTAACTGATAGTTATTATTGTTGGAAAAATATTACAGAAGG
ATTAGGTTTAGAAGGACCACAAAGTCCTTGGTGGCCTATAATACAAAATATACGAGAAAA
AGAGATAGTTTATTTTGCTTGGGTACCTGGTCACAAAGGGATATATGGTAATCAATTGGC
AGATGAAGCCGCAAAAATAAAAGAAGAAATCATGCTAGCATACCAAGGCACACAAATTAA
AGAGAAAAGAGATGAAGATGCAGGGTTTGACTTATGTGTTCCTTATGACATCATGATACC
TGTATCTGACACAAAAATCATACCCACAGATGTAAAAATTCAAGTTCCTCCTAATAGCTT
TGGATGGGTCACTGGGAAATCATCAATGGCAAAACAGGGGTTATTAATTAATGGAGGAAT
AATTGATGAAGGATATACAGGAGAAATACAAGTGATATGTACTAATATTGGAAAAAGTAA
TATTAAATTAATAGAGGGACAAAAATTTGCACAATTAATTATACTACAGCATCACTCAAA
TTCCAGACAGCCTTGGGATGAAAATAAAATATCTCAGAGAGGGGATAAAGGATTTGGAAG
TACAGGAGTATTCTGGGTAGAAAATATTCAGGAAGCACAAGATGAACATGAGAATTGGCA
TACATCACCAAAGATATTGGCAAGAAATTATAAGATACCATTGACTGTAGCAAAACAGAT
AACTCAAGAATGTCCTCATTGCACTAAGCAAGGATCAGGACCTGCAGGTTGTGTCATGAG
ATCTCCTAATCATTGGCAGGCAGATTGCACACATTTGGACAATAAGATAATATTGACTTT
TGTAGAGTCAAATTCAGGATACATACATGCTACATTATTGTCAAAAGAAAATGCATTATG
TACTTCATTGGCTATTTTAGAATGGGCAAGATTGTTTTCACCAAAGTCCTTACACACAGA
TAACGGCACTAATTTTGTGGCAGAACCAGTTGTAAATTTGTTGAAGTTCCTAAAGATAGC
ACATACCACAGGAATACCATATCATCCAGAAAGTCAGGGTATTGTAGAAAGGGCAAATAG
GACCTTGAAAGAGAAGATTCAAAGTCATAGAGACAACACTCAAACACTGGAGGCAGCTTT
ACAACTTGCTCTCATTACTTGTAACAAAGGGAGGGAAAGTATGGGAGGACAGACACCATG
GGAAGTATTTATCACTAATCAAGCACAAGTAATACATGAGAAACTTTTACTACAGCAAGC
ACAATCCTCCAAAAAATTTTGTTTTTACAAAATCCCTGGTGAACATGATTGGAAGGGACC
TACTAGGGTGCTGTGGAAGGGTGATGGTGCAGTAGTAGTTAATGATGAAGGAAAGGGAAT
AATTGCTGTACCATTAACCAGGACTAAGTTACTAATAAAACCAAATTGAGTATTGTTGCA
GGAAGCAAGACCCAACTACCATTGTCAGCTGTGTTTCCTGACCTCAATATTTGTTATAAG
GTTTGATATGAATCCCAGGGGGAATCTCAACCCCTATTACCCAACAGTCAGAAAAATCTA
AGTGTGAGGAGAACACAATGTTTCAACCTTATTGTTATAATAATGACAGTAAGAACAGCA
TGGCAGAATCGAAGGAAGCAAGAGACCAAGAAATGAACCTGAAAGAAGAATCTAAAGAAG
AAAAAAGAAGAAATGACTGGTGGAAAATAGGTATGTTTCTGTTATGCTTAGCAGGAACTA
CTGGAGGAATACTTTGGTGGTATGAAGGACTCCCACAGCAACATTATATAGGGTTGGTGG
CGATAGGGGGAAGATTAAACGGATCTGGCCAATCAAATGCTATAGAATGCTGGGGTTCCT
TCCCGGGGTGTAGACCATTTCAAAATTACTTCAGTTATGAGACCAATAGAAGCATGCATA
TGGATAATAATACTGCTACATTATTAGAAGCTTTAACCAATATAACTGCTCTATAAATAA
CAAAACAGAATTAGAAACATGGAAGTTAGTAAAGACTTCTGGCATAACTCCTTTACCTAT
TTCTTCTGAAGCTAACACTGGACTAATTAGACATAAGAGAGATTTTGGTATAAGTGCAAT
AGTGGCAGCTATTGTAGCCGCTACTGCTATTGCTGCTAGCGCTACTATGTCTTATGTTGC
TCTAACTGAGGTTAACAAAATAATGGAAGTACAAAATCATACTTTTGAGGTAGAAAATAG
TACTCTAAATGGTATGGATTTAATAGAACGACAAATAAAGATATTATATGCTATGATTCT
TCAAACACATGCAGATGTTCAACTGTTAAAGGAAAGACAACAGGTAGAGGAGACATTTAA
TTTAATTGGATGTATAGAAAGAACACATGTATTTTGTCATACTGGTCATCCCTGGAATAT
GTCATGGGGACATTTAAATGAGTCAACACAATGGGATGACTGGGTAAGCAAAATGGAAGA
TTTAAATCAAGAGATACTAACTACACTTCATGGAGCCAGGAACAATTTGGCACAATCCAT
GATAACATTCAATACACCAGATAGTATAGCTCAATTTGGAAAAGACCTTTGGAGTCATAT
TGGAAATTGGATTCCTGGATTGGGAGCTTCCATTATAAAATATATAGTGATGTTTTTGCT
TATTTATTTGTTACTAACCTCTTCGCCTAAGATCCTCAGGGCCCTCTGGAAGGTGACCAG
TGGTGCAGGGTCCTCCGGCAGTCGTTACCTGAAGAAAAAATTCCATCACAAACATGCATC
GCGAGAAGACACCTGGGACCAGGCCCAACACAACATACACCTAGCAGGCGTGACCGGTGG
ATCAGGGGACAAATACTACAAGCAGAAGTACTCCAGGAACGACTGGAATGGAGAATCAGA
GGAGTACAACAGGCGGCCAAAGAGCTGGGTGAAGTCAATCGAGGCATTTGGAGAGAGCTA
TATTTCCGAGAAGACCAAAGGGGAGATTTCTCAGCCTGGGGCGGCTATCAACGAGCACAA
GAACGGCTCTGGGGGGAACAATCCTCACCAAGGGTCCTTAGACCTGGAGATTCGAAGCGA
AGGAGGAAACATTTATGACTGTTGCATTAAAGCCCAAGAAGGAACTCTCGCTATCCCTTG
CTGTGGATTTCCCTTATGGCTATTTTGGGGGTCGACCCGGGCGGCCGCACTAGAGGAATT
CGCCCCTCTCCCTCCCCCCCCCCTAACGTTACTGGCCGAAGCCGCTTGGAATAAGGCCGG
TGTGTGTTTGTCTATATGTGATTTTCCACCATATTGCCGTCTTTTGGCAATGTGAGGGCC
CGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCTAGGGGTCTTTCCCCTCTCGCCAAA
GGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCAGTTCCTCTGGAAGCTTCTTGAAGA
CAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGGAACCCCCCACCTGGCGACAGGTGC
CTCTGCGGCCAAAAGCCACGTGTATAAGATACACCTGCAAAGGCGGCACAACCCCAGTGC

CACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAATGGCTCTCCTCAAGCGTAGTCAAC
AAGGGGCTGAAGGATGCCCAGAAGGTACCCCATTGTATGGGAATCTGATCTGGGGCCTCG
GTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAAAAGCTCTAGGCCCCCCGAACCAC
GGGGACGTGGTTTTCCTTTGAAAAACACGATGATAAGCTTGCCACAACCCCGTACCAAAG
ATGGATAGATCCGGAAAGCCTGAACTCACCGCGACGTCTGTCGAGAAGTTTCTGATCGAA
AAGTTCGACAGCGTCTCCGACCTGATGCAGCTCTCGGAGGGCGAAGAATCTCGTGCTTTC
AGCTTCGATGTAGGAGGGCGTGGATATGTCCTGCGGGTAAATAGCTGCGCCGATGGTTTC
TACAAAGATCGTTATGTTTATCGGCACTTTGCATCGGCCGCGCTCCCGATTCCGGAAGTG
CTTGACATTGGGGAATTCAGCGAGAGCCTGACCTATTGCATCTCCCGCCGTGCACAGGGT
GTCACGTTGCAAGACCTGCCTGAAACCGAACTGCCCGCTGTTCTGCAGCCGGTCGCGGAG
GCCATGGATGCGATCGCTGCGGCCGATCTTAGCCAGACGAGCGGGTTCGGCCCATTCGGA
CCGCAAGGAATCGGTCAATACACTACATGGCGTGATTTCATATGCGCGATTGCTGATCCC
CATGTGTATCACTGGCAAACTGTGATGGACGACACCGTCAGTGCGTCCGTCGCGCAGGCT
CTCGATGAGCTGATGCTTTGGGCCGAGGACTGCCCCGAAGTCCGGCACCTCGTGCACGCG
GATTTCGGCTCCAACAATGTCCTGACGGACAATGGCCGCATAACAGCGGTCATTGACTGG
AGCGAGGCGATGTTCGGGGATTCCCAATACGAGGTCGCCAACATCTTCTTCTGGAGGCCG
TGGTTGGCTTGTATGGAGCAGCAGACGCGCTACTTCGAGCGGAGGCATCCGGAGCTTGCA
GGATCGCCGCGGCTCCGGGCGTATATGCTCCGCATTGGTCTTGACCAACTCTATCAGAGC
TTGGTTGACGGCAATTTCGATGATGCAGCTTGGGCGCAGGGTCGATGCGACGCAATCGTC
CGATCCGGAGCCGGGACTGTCGGGCGTACACAAATCGCCCGCAGAAGCGCGGCCGTCTGG
ACCGATGGCTGTGTAGAAGTACTCGCCGATAGTGGAAACCGACGCCCCAGCACTCGTCCG
AGGGCAAAGGAATAGAGTAGATGCCGACCGAACAAGAGCTGATTTCGAGAACGCCTCAGC
CAGCAACTCGCGCGAGCCTAGCAAGGCAAATGCGAGAGAACGGCCTTACGCTTGGTGGCA
CAGTTCTCGTCCACAGTTCGCTAAGCTCGCTCGGCTGGGTCGCGGGAGGGCCGGTCGCAG
TGATTCAGGCCCTTCTGGATTGTGTTGGTCCCCAGGGCACGATTGTCATGCCCACGCACT
CGGGTGATCTGACTGATCCCGCAGATTGGAGATCGCCGCCCGTGCCTGCCGATTGGGTGC
AGATCTAGAGCTCGCTGATCAGCCTCGACTGTGCCTCTAGTTGCCAGCCATCTGTTGTTT
GCCCCTCCCCCGTGCCTTCCTTGACCCTGGAAGGTGCCACTCCCACTGTCCTTTCCTAAT
AAAATGAGGAAATTGCATCGCATTGTCTGAGTAGGTGTCATTCTATTCTGGGGGGTGGGG
TGGGGCAGGACAGCAAGGGGGAGGATTGGGAAGACAATAGCAGGCATGCTGGGGATGCGG
TGGGCTCTATGGCTTCTGAGGCGGAAAGAACCAGCTGGGGCTCGAGTGCATTCTAGTTGT
GGTTTGTCCAAACTCATCAATGTATCTTATCATGTCTGTATACCGTCGACCTCTAGCTAG
AGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATT
CCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGC
TAACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGC
CAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCT
TCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCA
GCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAAC
ATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTT
TTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGG
CGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGC
TCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGC
GTGGCGCTTTCTCAATGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCC
AAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAAC
TATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGT
AACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCT
AACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACC
TTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGT
TTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTG
ATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTC
ATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAA
TCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAG
GCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTG
TAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGA
GACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAG
CGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAA
GCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGC
ATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCA
AGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCG
ATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCAT
AATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACC
AAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGG
GATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCG
GGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGT
GCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACA
GGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATA
CTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATAC
ATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAA
GTGCCACCTGACGTCGACGGATCGGG

**pONY8ZA CMVHyb (SEQ ID No 52)**

AGATCTTGAATAATAAAATGTGTGTTTGTCCGAAATACGCGTTTTGAGATTTCTGTCGCC

GACTAAATTCATGTCGCGCGATAGTGGTGTTTATCGCCGATAGAGATGGCGATATTGGAA
AAATTGATATTTGAAAATATGGCATATTGAAAATGTCGCCGATGTGAGTTTCTGTGTAAC
TGATATCGCCATTTTTCCAAAAGTGATTTTTGGGCATACGCGATATCTGGCGATAGCGCT
TATATCGTTTACGGGGGATGGCGATAGACGACTTTGGTGACTTGGGCGATTCTGTGTGTC
GCAAATATCGCAGTTTCGATATAGGTGACAGACGATATGAGGCTATATCGCCGATAGAGG
CGACATCAAGCTGGCACATGGCCAATGCATATCGATCTATACATTGAATCAATATTGGCC
ATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCA
TACGTTGTATCCATATCGTAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCC
ATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCA
TAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACC
GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAAT
AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT
ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC
CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTA
CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGG
ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT
GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTGCGATCGCCCGCC
CCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGT
TTAGTGAACCGGGCACTCAGATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGG
CCTTTGTAATAAATATAATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATC
CTACAGTTGGCGCCCGAACAGGGACCTGAGAGGGGCGCAGACCCTACCTGTTGAACCTGG
CTGATCGTAGGATCCCCGGGACAGCAGAGGAGAACTTACAGAAGTCTTCTGGAGGTGTTC
CTGGCCAGAACACAGGAGGACAGGTAAGATTGGGAGACCCTTTGACATTGGAGCAAGGCG
CTCAAGAAGTTAGAGAAGGTGACGGTACAAGGGTCTCAGAAATTAACTACTGGTAACTGT
AATTGGGCGCTAAGTCTAGTAGACTTATTTCATTGATACCAACTTTGTAAAAGAAAAGGA
CTGGCAGCTGAGGGATTGTCATTCCATTGCTGGAAGATTGTAACTCAGACGCTGTCAGGA
CAAGAAAGAGAGGCCTTTGAAAGAACATTGGTGGGCAATTTCTGCTGTAAAGATTGGGCC
TCCAGATTAATAATTGTAGTAGATTGGAAAGGCATCATTCCAGCTCCTAAGAGCGAAATA
TTGAAAAGAAGACTGCTAATAAAAAGCAGTCTGAGCCCTCTGAAGAATATCTCTAGAACT
AGTGGATCCCCCGGGCTGCAGGAGTGGGGAGGCACGATGGCCGCTTTGGTCGAGGCGGAT
CCGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCC
ATTGCATACGTTGTATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATT
ACCGCCATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATT
AGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGG
CTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAAC
GCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTT
GGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAA
ATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTA
CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGG
GCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG
GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCC
ATTGACGCAAATGGGCGGTAGGCATGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT
AGTGAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAGACA
CCGGGACCGATCCAGCCTCCGCGGCCCCAAGCTTCAGCTGCTCGAGGATCTGCGGATCCG
GGGAATTCCCCAGTCTCAGGATCCACCATGGGGGATCCCGTCGTTTTACAACGTCGTGAC
TGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGC
TGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAAT
GGCGAATGGCGCTTTGCCTGGTTTCCGGCACCAGAAGCGGTGCCGGAAAGCTGGCTGGAG
TGCGATCTTCCTGAGGCCGATACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTAC
GATGCGCCCATCTACACCAACGTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCC
ACGGAGAATCCGACGGGTTGTTACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAG
GAAGGCCAGACGCGAATTATTTTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAAC
GGGCGCTGGGTCGGTTACGGCCAGGACAGTCGTTTGCCGTCTGAATTTGACCTGAGCGCA
TTTTTACGCGCCGGAGAAAACCGCCTCGCGGTGATGGTGCTGCGTTGGAGTGACGGCAGT
TATCTGGAAGATCAGGATATGTGGCGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCTG
CATAAACCGACTACACAAATCAGCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTC
AGCCGCGCTGTACTGGAGGCTGAAGTTCAGATGTGCGGCGAGTTGCGTGACTACCTACGG
GTAACAGTTTCTTTATGGCAGGGTGAAACGCAGGTCGCCAGCGGCACCGCGCCTTTCGGC
GGTGAAATTATCGATGAGCGTGGTGGTTATGCCGATCGCGTCACACTACGTCTGAACGTC
GAAAACCCGAAACTGTGGAGCGCCGAAATCCCGAATCTCTATCGTGCGGTGGTTGAACTG
CACACCGCCGACGGCACGCTGATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTG
CGGATTGAAAATGGTCTGCTGCTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAAC
CGTCACGAGCATCATCCTCTGCATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGAT
ATCCTGCTGATGAAGCAGAACAACTTTAACGCCGTGCGCTGTTCGCATTATCCGAACCAT
CCGCTGTGGTACACGCTGTGCGACCGCTACGGCCTGTATGTGGTGGATGAAGCCAATATT
GAAACCCACGGCATGGTGCCAATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCG
ATGAGCGAACGCGTAACGCGAATGGTGCAGCGCGATCGTAATCACCCGAGTGTGATCATC
TGGTCGCTGGGGAATGAATCAGGCCACGGCGCTAATCACGACGCGCTGTATCGCTGGATC
AAATCTGTCGATCCTTCCCGCCCGGTGCAGTATGAAGGCGGCGGAGCCGACACCACGGCC
ACCGATATTATTTGCCCGATGTACGCGCGCGTGGATGAAGACCAGCCCTTCCCGGCTGTG
CCGAAATGGTCCATCAAAAAATGGCTTTCGCTACCTGGAGAGACGCGCCCGCTGATCCTT
TGCGAATACGCCCACGCGATGGGTAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGCG
TTTCGTCAGTATCCCCGTTTACAGGGCGGCTTCGTCTGGGACTGGGTGGATCAGTCGCTG
ATTAAATATGATGAAAACGGCAACCCGTGGTCGGCTTACGGCGGTGATTTTGGCGATACG

CCGAACGATCGCCAGTTCTGTATGAACGGTCTGGTCTTTGCCGACCGCACGCCGCATCCA
GCGCTGACGGAAGCAAAACACCAGCAGCAGTTTTTCCAGTTCCGTTTATCCGGGCAAACC
ATCGAAGTGACCAGCGAATACCTGTTCCGTCATAGCGATAACGAGCTCCTGCACTGGATG
GTGGCGCTGGATGGTAAGCCGCTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAA
GGTAAACAGTTGATTGAACTGCCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGG
CTCACAGTACGCGTAGTGCAACCGAACGCGACCGCATGGTCAGAAGCCGGGCACATCAGC
GCCTGGCAGCAGTGGCGTCTGGCGGAAAACCTCAGTGTGACGCTCCCCGCCGCGTCCCAC
GCCATCCCGCATCTGACCACCAGCGAAATGGATTTTTGCATCGAGCTGGGTAATAAGCGT
TGGCAATTTAACCGCCAGTCAGGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAA
CTGCTGACGCCGCTGCGCGATCAGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTA
AGTGAAGCGACCCGCATTGACCCTAACGCCTGGGTCGAACGCTGGAAGGCGGCGGGCCAT
TACCAGGCCGAAGCAGCGTTGTTGCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTG
ATTACGACCGCTCACGCGTGGCAGCATCAGGGGAAAACCTTATTTATCAGCCGGAAAACC
TACCGGATTGATGGTAGTGGTCAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGAT
ACACCGCATCCGGCGCGGATTGGCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTA
AACTGGCTCGGATTAGGGCCGCAAGAAAACTATCCCGACCGCCTTACTGCCGCCTGTTTT
GACCGCTGGGATCTGCCATTGTCAGACATGTATACCCCGTACGTCTTCCCGAGCGAAAAC
GGTCTGCGCTGCGGGACGCGCGAATTGAATTATGGCCCACACCAGTGGCGCGGCGACTTC
CAGTTCAACATCAGCCGCTACAGTCAACAGCAACTGATGGAAACCAGCCATCGCCATCTG
CTGCACGCGGAAGAAGGCACATGGCTGAATATCGACGGTTTCCATATGGGGATTGGTGGC
GACGACTCCTGGAGCCCGTCAGTATCGGCGGAATTCCAGCTGAGCGCCGGTCGCTACCAT
TACCAGTTGGTCTGGTGTCAAAAATAATAATAACCGGGCAGGGGGGATCCGCAGATCCGG
CTGTGGAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGT
ATGCAAAGCATGCCTGCAGGAATTCGATATCAAGCTTATCGATACCGTCGACCTCGAGGG
GGGGCCCGGTACCCAGCTTTTGTTCCCTTTAGTGAGGGTTAATTGCGCGGGAAGTATTTA
TCACTAATCAAGCACAAGTAATACATGAGAAACTTTTACTACAGCAAGCACAATCCTCCA
AAAAATTTTGTTTTTACAAAATCCCTGGTGAACATGATTGGAAGGGACCTACTAGGGTGC
TGTGGAAGGGTGATGGTGCAGTAGTAGTTAATGATGAAGGAAAGGGAATAATTGCTGTAC
CATTAACCAGGACTAAGTTACTAATAAAACCAAATTGAGTATTGTTGCAGGAAGCAAGAC
CCAACTACCATTGTCAGCTGTGTTTCCTGACCTCAATATTTGTTATAAGGTTTGATATGA
ATCCCAGGGGGAATCTCAACCCCTATTACCCAACAGTCAGAAAAATCTAAGTGTGAGGAG
AACACAATGTTTCAACCTTATTGTTATAATAATGACAGTAAGAACAGCATGGCAGAATCG
AAGGAAGCAAGAGACCAAGAATGAACCTGAAAGAAGAATCTAAAGAAGAAAAAAGAAGAA
ATGACTGGTGGAAAATAGGTATGTTTCTGTTATGCTTAGCAGGAACTACTGGAGGAATAC
TTTGGTGGTATGAAGGACTCCCACAGCAACATTATATAGGGTTGGTGGCGATAGGGGGAA
GATTAAACGGATCTGGCCAATCAAATGCTATAGAATGCTGGGGTTCCTTCCCGGGGTGTA
GACCATTTCAAAATTACTTCAGTTATGAGACCAATAGAAGCATGCATATGGATAATAATA
CTGCTACATTATTAGAAGCTTTAACCAATATAACTGCTCTATAAATAACAAAACAGAATT
AGAAACATGGAAGTTAGTAAAGACTTCTGGCATAACTCCTTTACCTATTTCTTCTGAAGC
TAACACTGGACTAATTAGACATAAGAGAGATTTTGGTATAAGTGCAATAGTGGCAGCTAT
TGTAGCCGCTACTGCTATTGCTGCTAGCGCTACTATGTCTTATGTTGCTCTAACTGAGGT
TAACAAAATAATGGAAGTACAAAATCATACTTTTGAGGTAGAAAATAGTACTCTAAATGG
TATGGATTTAATAGAACGACAAATAAAGATATTATATGCTATGATTCTTCAAACACATGC
AGATGTTCAACTGTTAAAGGAAAGACAACAGGTAGAGGAGACATTTAATTTAATTGGATG
TATAGAAAGAACACATGTATTTTGTCATACTGGTCATCCCTGGAATATGTCATGGGGACA
TTTAAATGAGTCAACACAATGGGATGACTGGGTAAGCAAAATGGAAGATTTAAATCAAGA
GATACTAACTACACTTCATGGAGCCAGGAACAATTTGGCACAATCCATGATAACATTCAA
TACACCAGATAGTATAGCTCAATTTGGAAAAGACCTTTGGAGTCATATTGGAAATTGGAT
TCCTGGATTGGGAGCTTCCATTATAAAATATATAGTGATGTTTTTGCTTATTTATTTGTT
ACTAACCTCTTCGCCTAAGATCCTCAGGGCCCTCTGGAAGGTGACCAGTGGTGCAGGGTC
CTCCGGCAGTCGTTACCTGAAGAAAAAATTCCATCACAAACATGCATCGCGAGAAGACAC
CTGGGACCAGGCCCAACACAACATACACCTAGCAGGCGTGACCGGTGGATCAGGGGACAA
ATACTACAAGCAGAAGTACTCCAGGAACGACTGGAATGGAGAATCAGAGGAGTACAACAG
GCGGCCAAAGAGCTGGGTGAAGTCAATCGAGGCATTTGGAGAGAGCTATATTTCCGAGAA
GACCAAAGGGGAGATTTCTCAGCCTGGGGCGGCTATCAACGAGCACAAGAACGGCTCTGG
GGGGAACAATCCTCACCAAGGGTCCTTAGACCTGGAGATTCGAAGCGAAGGAGGAAACAT
TTATGACTGTTGCATTAAAGCCCAAGAAGGAACTCTCGCTATCCCTTGCTGTGGATTTCC
CTTATGGCTATTTTGGGGACTAGTAATTATAGTAGGACGCATAGCAGGCTATGGATTACG
TGGACTCGCTGTTATAATAAGGATTTGTATTAGAGGCTTAAATTTGATATTTGAAATAAT
CAGAAAAATGCTTGATTATATTGGAAGAGCTTTAAATCCTGGCACATCTCATGTATCAAT
GCCTCAGTATGTTTAGAAAAACAAGGGGGGAACTGTGGGGTTTTTATGAGGGGTTTTATA
AACTGCAGGAGTGGGGAGGCACGATGGCCGCTTTGGTCGAGGCGGATCCGGCCATTAGCC
ATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTG
TATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCATGTTGA
CATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCA
TATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAAC
GACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACT
TTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA
GTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGG
CATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTA
GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGG
TTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGG
CACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATG
GGCGGTAGGCATGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGGGCA

CTCAGATTCTGCGGTCTGAGTCCCTTCTCTGCTGGGCTGAAAAGGCCTTTGTAATAAATA
TAATTCTCTACTCAGTCCCTGTCTCTAGTTTGTCTGTTCGAGATCCTACAGAGCTCATGC
CTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACAATTCC
ACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTA
ACTCACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCA
GCTGCATTAATGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTC
CGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGC
TCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACAT
GTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTT
CCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCG
AAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTC
TCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGT
GGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAA
GCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTA
TCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAA
CAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAA
CTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTT
CGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT
TTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGAT
CTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCAT
GAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATC
AATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGC
ACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA
GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGA
CCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCG
CAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGC
TAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTACAGGCAT
CGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAG
GCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGAT
CGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAA
TTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAA
GTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGA
TAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGG
GCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGC
ACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGG
AAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACT
CTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACAT
ATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGT
GCCACCTAAATTGTAAGCGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATC
AGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAG
ACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTG
GACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCA
TCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACCCTAAA
GGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCAACCTGGCTTATCGAAATTAATACG
ACTCACTATAGG

## PEsynGP (SEQ ID No 53)

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTA
TTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCC
AATATGACCGCCATGTTGGCATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGG
GTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCC
GCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCAT
AGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGA
CGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACAC
CAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGT
CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTG
CGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA
AGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCAC
AGTTAAATTGCTAACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGT
GACTCTCTTAAGGTAGCCTTGCAGAAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAA
GGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCTTGTCGAGACAGAGAAGACT
CTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCAC
AGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACT
ATAGGCTAGAGAATTCGCCACCATGGGCGATCCCCTCACCTGGTCCAAAGCCCTGAAGAA
ACTGGAAAAAGTCACCGTTCAGGGTAGCCAAAAGCTTACCACAGGCAATTGCAACTGGGC
ATTGTCCCTGGTGGATCTTTTCCACGACACTAATTTCGTTAAGGAGAAAGATTGGCAACT
CAGAGACGTGATCCCCCTCTTGGAGGACGTGACCCAAACATTGTCTGGGCAGGAGCGCGA
AGCTTTCGAGCGCACCTGGTGGGCCATCAGCGCAGTCAAAATGGGGCTGCAAATCAACAA
CGTGGTTGACGGTAAAGCTAGCTTTCAACTGCTCCGCGCTAAGTACGAGAAGAAAACCGC
CAACAAGAAACAATCCGAACCTAGCGAGGAGTACCCAATTATGATCGACGGCGCCGGCAA
TAGGAACTTCCGCCCACTGACTCCCAGGGGCTATACCACCTGGGTCAACACCATCCAGAC
AAACGGACTTTTGAACGAAGCCTCCCAGAACCTGTTCGGCATCCTGTCTGTGGACTGCAC
CTCCGAAGAAATGAATGCTTTTCTCGACGTGGTGCCAGGACAGGCTGGACAGAAACAGAT
CCTGCTCGATGCCATTGACAAGATCGCCGACGACTGGGATAATCGCCACCCCCTGCCAAA
CGCCCCTCTGGTGGCTCCCCCACAGGGGCCTATCCCTATGACCGCTAGGTTCATTAGGGG
ACTGGGGGTGCCCCGCGAACGCCAGATGGAGCCAGCATTTGACCAATTTAGGCAGACCTA
CAGACAGTGGATCATCGAAGCCATGAGCGAGGGGATTAAAGTCATGATCGGAAAGCCCAA
GGCACAGAACATCAGGCAGGGGGCCAAGGAACCATACCCTGAGTTTGTCGACAGGCTTCT
GTCCCAGATTAAATCCGAAGGCCACCCTCAGGAGATCTCCAAGTTCTTGACAGACACACT
GACTATCCAAAATGCAAATGAAGAGTGCAGAAACGCCATGAGGCACCTCAGACCTGAAGA
TACCCTGGAGGAGAAAATGTACGCATGTCGCGACATTGGCACTACCAAGCAAAAGATGAT
GCTGCTCGCCAAGGCTCTGCAAACCGGCCTGGCTGGTCCATTCAAAGGAGGAGCACTGAA
GGGAGGTCCATTGAAAGCTGCACAAACATGTTATAATTGTGGGAAGCCAGGACATTTATC
TAGTCAATGTAGAGCACCTAAAGTCTGTTTTAAATGTAAACAGCCTGGACATTTCTCAAA
GCAATGCAGAAGTGTTCCAAAAAACGGGAAGCAAGGGGCTCAAGGGAGGCCCCAGAAACA
AACTTTCCCGATACAACAGAAGAGTCAGCACAACAAATCTGTTGTACAAGAGACTCCTCA
GACTCAAAATCTGTACCCAGATCTGAGCGAAATAAAAAAGGAATACAATGTCAAGGAGAA
GGATCAAGTAGAGGATCTCAACCTGGACAGTTTGTGGGAGTAACATACAATCTCGAGAAG
AGGCCCACTACCATCGTCCTGATCAATGACACCCCTCTTAATGTGCTGCTGGACACCGGA
GCCGACACCAGCGTTCTCACTACTGCTCACTATAACAGACTGAAATACAGAGGAAGGAAA
TACCAGGGCACAGGCATCATCGGCGTTGGAGGCAACGTCGAAACCTTTTCCACTCCTGTC
ACCATCAAAAAGAAGGGGAGACACATTAAAACCAGAATGCTGGTCGCCGACATCCCCGTC
ACCATCCTTGGCAGAGACATTCTCCAGGACCTGGGCGCTAAACTCGTGCTGGCACAACTG
TCTAAGGAAATCAAGTTCCGCAAGATCGAGCTGAAAGAGGGCACAATGGGTCCAAAAATC
CCCCAGTGGCCCCTGACCAAAGAGAAGCTTGAGGGCGCTAAGGAAATCGTGCAGCGCCTG
CTTTCTGAGGGCAAGATTAGCGAGGCCAGCGACAATAACCCTTACAACAGCCCCATCTTT

GTGATTAAGAAAAGGAGCGGCAAATGGAGACTCCTGCAGGACCTGAGGGAACTCAACAAG
ACCGTCCAGGTCGGAACTGAGATCTCTCGCGGACTGCCTCACCCCGGCGGCCTGATTAAA
TGCAAGCACATGACAGTCCTTGACATTGGAGACGCTTATTTTACCATCCCCCTCGATCCT
GAATTTCGCCCCTATACTGCTTTTACCATCCCCAGCATCAATCACCAGGAGCCCGATAAA
CGCTATGTGTGGAAGTGCCTCCCCCAGGGATTTGTGCTTAGCCCCTACATTTACCAGAAG
ACACTTCAAGAGATCCTCCAACCTTTCCGCGAAAGATACCCAGAGGTTCAACTCTACCAA
TATATGGACGACCTGTTCATGGGGTCCAACGGGTCTAAGAAGCAGCACAAGGAACTCATC
ATCGAACTGAGGGCAATCCTCCTGGAGAAAGGCTTCGAGACACCCGACGACAAGCTGCAA
GAAGTTCCTCCATATAGCTGGCTGGGCTACCAGCTTTGCCCTGAAAACTGGAAAGTCCAG
AAGATGCAGTTGGATATGGTCAAGAACCCAACACTGAACGACGTCCAGAAGCTCATGGGC
AATATTACCTGGATGAGCTCCGGAATCCCTGGGCTTACCGTTAAGCACATTGCCGCAACT
ACAAAAGGATGCCTGGAGTTGAACCAGAAGGTCATTTGGACAGAGGAAGCTCAGAAGGAA
CTGGAGGAGAATAATGAAAAGATTAAGAATGCTCAAGGGCTCCAATACTACAATCCCGAA
GAAGAAATGTTGTGCGAGGTCGAAATCACTAAGAACTACGAAGCCACCTATGTCATCAAA
CAGTCCCAAGGCATCTTGTGGGCCGGAAAGAAAATCATGAAGGCCAACAAAGGCTGGTCC
ACCGTTAAAAATCTGATGCTCCTGCTCCAGCACGTCGCCACCGAGTCTATCACCCGCGTC
GGCAAGTGCCCCACCTTCAAAGTTCCCTTCACTAAGGAGCAGGTGATGTGGGAGATGCAA
AAAGGCTGGTACTACTCTTGGCTTCCCGAGATCGTCTACACCCACCAAGTGGTGCACGAC
GACTGGAGAATGAAGCTTGTCGAGGAGCCCACTAGCGGAATTACAATCTATACCGACGGC
GGAAAGCAAAACGGAGAGGGAATCGCTGCATACGTCACATCTAACGGCCGCACCAAGCAA
AAGAGGCTCGGCCCTGTCACTCACCAGGTGGCTGAGAGGATGGCTATCCAGATGGCCCTT
GAGGACACTAGAGACAAGCAGGTGAACATTGTGACTGACAGCTACTACTGCTGGAAAAAC
ATCACAGAGGGCCTTGGCCTGGAGGGACCCCAGTCTCCCTGGTGGCCTATCATCCAGAAT
ATCCGCGAAAAGGAAATTGTCTATTTCGCCTGGGTGCCTGGACACAAAGGAATTTACGGC
AACCAACTCGCCGATGAAGCCGCCAAAATTAAAGAGGAAATCATGCTTGCCTACCAGGGC
ACACAGATTAAGGAGAAGAGAGACGAGGACGCTGGCTTTGACCTGTGTGTGCCATACGAC
ATCATGATTCCCGTTAGCGACACAAAGATCATTCCAACCGATGTCAAGATCCAGGTGCCA
CCCAATTCATTTGGTTGGGTGACCGGAAAGTCCAGCATGGCTAAGCAGGGTCTTCTGATT
AACGGGGGAATCATTGATGAAGGATACACCGGCGAAATCCAGGTGATCTGCACAAATATC
GGCAAAAGCAATATTAAGCTTATCGAAGGGCAGAAGTTCGCTCAACTCATCATCCTCCAG
CACCACAGCAATTCAAGACAACCTTGGGACGAAAACAAGATTAGCCAGAGAGGTGACAAG
GGCTTCGGCAGCACAGGTGTGTTCTGGGTGGAGAACATCCAGGAAGCACAGGACGAGCAC
GAGAATTGGCACACCTCCCCTAAGATTTTGGCCCGCAATTACAAGATCCCACTGACTGTG
GCTAAGCAGATCACACAGGAATGCCCCCACTGCACCAAACAAGGTTCTGGCCCCGCCGGC
TGCGTGATGAGGTCCCCCAATCACTGGCAGGCAGATTGCACCCACCTCGACAACAAAATT
ATCCTGACCTTCGTGGAGAGCAATTCCGGCTACATCCACGCAACACTCCTCTCCAAGGAA
AATGCATTGTGCACCTCCCTCGCAATTCTGGAATGGGCCAGGCTGTTCTCTCCAAAATCC
CTGCACACCGACAACGGCACCAACTTTGTGGCTGAACCTGTGGTGAATCTGCTGAAGTTC
CTGAAAATCGCCCACACCACTGGCATTCCCTATCACCCTGAAAGCCAGGGCATTGTCGAG
AGGGCCAACAGAACTCTGAAAGAAAAGATCCAATCTCACAGAGACAATACACAGACATTG
GAGGCCGCACTTCAGCTCGCCCTTATCACCTGCAACAAAGGAAGAGAAAGCATGGGCGGC
CAGACCCCCTGGGAGGTCTTCATCACTAACCAGGCCCAGGTCATCCATGAAAAGCTGCTC
TTGCAGCAGGCCCAGTCCTCCAAAAAGTTCTGCTTTTATAAGATCCCCGGTGAGCACGAC
TGGAAAGGTCCTACAAGAGTTTTGTGGAAAGGAGACGGCGCAGTTGTGGTGAACGATGAG
GGCAAGGGGATCATCGCTGTGCCCCTGACACGCACCAAGCTTCTCATCAAGCCAAACTGA
ACCCGGGGCGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATA
CATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGTGA
AATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAACAA
CAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAAAG
CAAGTAAAACCTCTACAAATGTGGTAAAATCCGATAAGGATCGATCCGGGCTGGCGTAAT
AGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGG
ACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCG
CTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCA
CGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTA
GAGCTTTACGGCACCTCGACCGCAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGGGC
CATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTG
GACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTAT
AAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAATATTTA
ACGCGAATTTTAACAAAATATTAACGTTTACAATTTCGCCTGATGCGGTATTTTCTCCTT
ACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTGCGCAGCACCATGGCCTGAAA
TAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGA
ATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAA
GCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCA
GAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGC
CCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTT
TTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAG

GAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTGATTCTTCTGACACAACAGTCT
CGAACTTAAGGCTAGAGCCACCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCGGC
CGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCTGA
TGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGACCT
GTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACGAC
GGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTGCT
ATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAAGT
ATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATT
CGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTTGT
CGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCCAG
GCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGCTT
GCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTGGG
TGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGG
CGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAGCG
CATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAATG
ACCGACCAAGCGACGCCCAACCTGCCATCACGATGGCCGCAATAAAATATCTTTATTTTC
ATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGCGATAAGGATCCGCGTATGGTG
CACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCAAC
ACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCTGT
GACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCGAG
ACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTTTC
TTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTT
CTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATA
ATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTT
TGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGC
TGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAGAT
CCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTGCT
ATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATACA
CTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGATGG
CATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCCAA
CTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATGGG
GGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAACGA
CGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACTGG
CGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAAGT
TGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCTGG
AGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCCTC
CCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGACA
GATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTACTC
ATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAGAT
CCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCGTC
AGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATCTG
CTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAGCT
ACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTCCT
TCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATACCT
CGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACCGG
GTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGTTC
GTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGTGA
GCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGCGG
CAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTTTA
TAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCAGG
GGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTTTG
CTGGCCTTTTGCTCACATGGCTCGACAGATCT

**PESDSYNGP (SEQ ID No 54)**

TCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTA
TTGGCCATTGCATACGTTGTATCTATATCATAATATGTACATTTATATTGGCTCATGTCC
AATATGACCGCCATGTTGGCATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGG
GTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATAACTTACGGTAAATGGCCC
GCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCAT
AGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC
CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGA
CGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTG
GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACAC
CAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGT

CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTG
CGATCGCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA
AGCAGAGCTCGTTTAGTGAACCGTCAGATCACTAGAAGCTTTATTGCGGTAGTTTATCAC
AGTTAAATTGCTAACGCAGTCAGTGCTTCTGACACAACAGTCTCGAACTTAAGCTGCAGT
GACTCTCTTAAGGTAGCCTTGCAGAAGTTGGTCGTGAGGCACTGGGCAGGTAAGTATCAA
GGTTACAAGACAGGTTTAAGGAGACCAATAGAAACTGGGCTTGTCGAGACAGAGAAGACT
CTTGCGTTTCTGATAGGCACCTATTGGTCTTACTGACATCCACTTTGCCTTTCTCTCCAC
AGGTGTCCACTCCCAGTTCAATTACAGCTCTTAAGGCTAGAGTACTTAATACGACTCACT
ATAGGCTAGAGAATTCCAGGTAAGATGGGCGATCCCCTCACCTGGTCCAAAGCCCTGAAG
AAACTGGAAAAAGTCACCGTTCAGGGTAGCCAAAAGCTTACCACAGGCAATTGCAACTGG
GCATTGTCCCTGGTGGATCTTTTCCACGACACTAATTTCGTTAAGGAGAAAGATTGGCAA
CTCAGAGACGTGATCCCCCTCTTGGAGGACGTGACCCAAACATTGTCTGGGCAGGAGCGC
GAAGCTTTCGAGCGCACCTGGTGGGCCATCAGCGCAGTCAAAATGGGGCTGCAAATCAAC
AACGTGGTTGACGGTAAAGCTAGCTTTCAACTGCTCCGCGCTAAGTACGAGAAGAAAACC
GCCAACAAGAAACAATCCGAACCTAGCGAGGAGTACCCAATTATGATCGACGGCGCCGGC
AATAGGAACTTCCGCCCACTGACTCCCAGGGGCTATACCACCTGGGTCAACACCATCCAG
ACAAACGGACTTTTGAACGAAGCCTCCCAGAACCTGTTCGGCATCCTGTCTGTGGACTGC
ACCTCCGAAGAAATGAATGCTTTTCTCGACGTGGTGCCAGGACAGGCTGGACAGAAACAG
ATCCTGCTCGATGCCATTGACAAGATCGCCGACGACTGGGATAATCGCCACCCCCTGCCA
AACGCCCCTCTGGTGGCTCCCCCACAGGGGCCTATCCCTATGACCGCTAGGTTCATTAGG
GGACTGGGGGTGCCCCGCGAACGCCAGATGGAGCCAGCATTTGACCAATTTAGGCAGACC
TACAGACAGTGGATCATCGAAGCCATGAGCGAGGGGATTAAAGTCATGATCGGAAAGCCC
AAGGCACAGAACATCAGGCAGGGGGCCAAGGAACCATACCCTGAGTTTGTCGACAGGCTT
CTGTCCCAGATTAAATCCGAAGGCCACCCTCAGGAGATCTCCAAGTTCTTGACAGACACA
CTGACTATCCAAAATGCAAATGAAGAGTGCAGAAACGCCATGAGGCACCTCAGACCTGAA
GATACCCTGGAGGAGAAAATGTACGCATGTCGCGACATTGGCACTACCAAGCAAAAGATG
ATGCTGCTCGCCAAGGCTCTGCAAACCGGCCTGGCTGGTCCATTCAAAGGAGGAGCACTG
AAGGGAGGTCCATTGAAAGCTGCACAAACATGTTATAATTGTGGGAAGCCAGGACATTTA
TCTAGTCAATGTAGAGCACCTAAAGTCTGTTTTAAATGTAAACAGCCTGGACATTTCTCA
AAGCAATGCAGAAGTGTTCCAAAAAACGGGAAGCAAGGGGCTCAAGGGAGGCCCCAGAAA
CAAACTTTCCCGATACAACAGAAGAGTCAGCACAACAAATCTGTTGTACAAGAGACTCCT
CAGACTCAAAATCTGTACCCAGATCTGAGCGAAATAAAAAAGGAATACAATGTCAAGGAG
AAGGATCAAGTAGAGGATCTCAACCTGGACAGTTTGTGGGAGTAACATACAATCTCGAGA
AGAGGCCCACTACCATCGTCCTGATCAATGACACCCCTCTTAATGTGCTGCTGGACACCG
GAGCCGACACCAGCGTTCTCACTACTGCTCACTATAACAGACTGAAATACAGAGGAAGGA
AATACCAGGGCACAGGCATCATCGGCGTTGGAGGCAACGTCGAAACCTTTTCCACTCCTG
TCACCATCAAAAAGAAGGGGAGACACATTAAAACCAGAATGCTGGTCGCCGACATCCCCG
TCACCATCCTTGGCAGAGACATTCTCCAGGACCTGGGCGCTAAACTCGTGCTGGCACAAC
TGTCTAAGGAAATCAAGTTCCGCAAGATCGAGCTGAAAGAGGGCACAATGGGTCCAAAAA
TCCCCCAGTGGCCCCTGACCAAAGAGAAGCTTGAGGGCGCTAAGGAAATCGTGCAGCGCC
TGCTTTCTGAGGGCAAGATTAGCGAGGCCAGCGACAATAACCCTTACAACAGCCCCATCT
TTGTGATTAAGAAAAGGAGCGGCAAATGGAGACTCCTGCAGGACCTGAGGGAACTCAACA
AGACCGTCCAGGTCGGAACTGAGATCTCTCGCGGACTGCCTCACCCCGGCGGCCTGATTA
AATGCAAGCACATGACAGTCCTTGACATTGGAGACGCTTATTTTACCATCCCCCTCGATC
CTGAATTTCGCCCCTATACTGCTTTTACCATCCCCAGCATCAATCACCAGGAGCCCGATA
AACGCTATGTGTGGAAGTGCCTCCCCCAGGGATTTGTGCTTAGCCCCTACATTTACCAGA
AGACACTTCAAGAGATCCTCCAACCTTTCCGCGAAAGATACCCAGAGGTTCAACTCTACC
AATATATGGACGACCTGTTCATGGGGTCCAACGGGTCTAAGAAGCAGCACAAGGAACTCA
TCATCGAACTGAGGGCAATCCTCCTGGAGAAAGGCTTCGAGACACCCGACGACAAGCTGC
AAGAAGTTCCTCCATATAGCTGGCTGGGCTACCAGCTTTGCCCTGAAAACTGGAAAGTCC
AGAAGATGCAGTTGGATATGGTCAAGAACCCAACACTGAACGACGTCCAGAAGCTCATGG
GCAATATTACCTGGATGAGCTCCGGAATCCCTGGGCTTACCGTTAAGCACATTGCCGCAA
CTACAAAAGGATGCCTGGAGTTGAACCAGAAGGTCATTTGGACAGAGGAAGCTCAGAAGG
AACTGGAGGAGAATAATGAAAAGATTAAGAATGCTCAAGGGCTCCAATACTACAATCCCG
AAGAAGAAATGTTGTGCGAGGTCGAAATCACTAAGAACTACGAAGCCACCTATGTCATCA
AACAGTCCCAAGGCATCTTGTGGGCCGGAAAGAAAATCATGAAGGCCAACAAAGGCTGGT
CCACCGTTAAAAATCTGATGCTCCTGCTCCAGCACGTCGCCACCGAGTCTATCACCCGCG
TCGGCAAGTGCCCCACCTTCAAAGTTCCCTTCACTAAGGAGCAGGTGATGTGGGAGATGC
AAAAAGGCTGGTACTACTCTTGGCTTCCCGAGATCGTCTACACCCACCAAGTGGTGCACG
ACGACTGGAGAATGAAGCTTGTCGAGGAGCCCACTAGCGGAATTACAATCTATACCGACG
GCGGAAAGCAAAACGGAGAGGGAATCGCTGCATACGTCACATCTAACGGCCGCACCAAGC
AAAAGAGGCTCGGCCCTGTCACTCACCAGGTGGCTGAGAGGATGGCTATCCAGATGGCCC
TTGAGGACACTAGAGACAAGCAGGTGAACATTGTGACTGACAGCTACTACTGCTGGAAAA
ACATCACAGAGGGCCTTGGCCTGGAGGGACCCCAGTCTCCCTGGTGGCCTATCATCCAGA
ATATCCGCGAAAAGGAAATTGTCTATTTCGCCTGGGTGCCTGGACACAAAGGAATTTACG
GCAACCAACTCGCCGATGAAGCCGCCAAAATTAAAGAGGAAATCATGCTTGCCTACCAGG

GCACACAGATTAAGGAGAAGAGAGACGAGGACGCTGGCTTTGACCTGTGTGTGCCATACG
ACATCATGATTCCCGTTAGCGACACAAAGATCATTCCAACCGATGTCAAGATCCAGGTGC
CACCCAATTCATTTGGTTGGGTGACCGGAAAGTCCAGCATGGCTAAGCAGGGTCTTCTGA
TTAACGGGGGAATCATTGATGAAGGATACACCGGCGAAATCCAGGTGATCTGCACAAATA
TCGGCAAAAGCAATATTAAGCTTATCGAAGGGCAGAAGTTCGCTCAACTCATCATCCTCC
AGCACCACAGCAATTCAAGACAACCTTGGGACGAAAACAAGATTAGCCAGAGAGGTGACA
AGGGCTTCGGCAGCACAGGTGTGTTCTGGGTGGAGAACATCCAGGAAGCACAGGACGAGC
ACGAGAATTGGCACACCTCCCCTAAGATTTTGGCCCGCAATTACAAGATCCCACTGACTG
TGGCTAAGCAGATCACACAGGAATGCCCCCACTGCACCAAACAAGGTTCTGGCCCCGCCG
GCTGCGTGATGAGGTCCCCCAATCACTGGCAGGCAGATTGCACCCACCTCGACAACAAAA
TTATCCTGACCTTCGTGGAGAGCAATTCCGGCTACATCCACGCAACACTCCTCTCCAAGG
AAAATGCATTGTGCACCTCCCTCGCAATTCTGGAATGGGCCAGGCTGTTCTCTCCAAAAT
CCCTGCACACCGACAACGGCACCAACTTTGTGGCTGAACCTGTGGTGAATCTGCTGAAGT
TCCTGAAAATCGCCCACACCACTGGCATTCCCTATCACCCTGAAAGCCAGGGCATTGTCG
AGAGGGCCAACAGAACTCTGAAAGAAAAGATCCAATCTCACAGAGACAATACACAGACAT
TGGAGGCCGCACTTCAGCTCGCCCTTATCACCTGCAACAAAGGAAGAGAAAGCATGGGCG
GCCAGACCCCCTGGGAGGTCTTCATCACTAACCAGGCCCAGGTCATCCATGAAAAGCTGC
TCTTGCAGCAGGCCCAGTCCTCCAAAAAGTTCTGCTTTTATAAGATCCCCGGTGAGCACG
ACTGGAAAGGTCCTACAAGAGTTTTGTGGAAAGGAGACGGCGCAGTTGTGGTGAACGATG
AGGGCAAGGGGATCATCGCTGTGCCCCTGACACGCACCAAGCTTCTCATCAAGCCAAACT
GAACCCGGGGCGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGA
TACATTGATGAGTTTGGACAAACCACAACTAGAATGCAGTGAAAAAAATGCTTTATTTGT
GAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTAAC
AACAACAATTGCATTCATTTTATGTTTCAGGTTCAGGGGGAGATGTGGGAGGTTTTTTAA
AGCAAGTAAAACCTCTACAAATGTGGTAAAATCCGATAAGGATCGATCCGGGCTGGCGTA
ATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAAT
GGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGAC
CGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGC
CACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATT
TAGAGCTTTACGGCACCTCGACCGCAAAAAACTTGATTTGGGTGATGGTTCACGTAGTGG
GCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAG
TGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTT
ATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAATATT
TAACGCGAATTTTAACAAAATATTAACGTTTACAATTTCGCCTGATGCGGTATTTTCTCC
TTACGCATCTGTGCGGTATTTCACACCGCATACGCGGATCTGCGCAGCACCATGGCCTGA
AATAACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTG
GAATGTGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCA
AAGCATGCATCTCAATTAGTCAGCAACCAGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGG
CAGAAGTATGCAAAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCC
GCCCATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAAT
TTTTTTTATTTATGCAGAGGCCGAGGCCGCCTCGGCCTCTGAGCTATTCCAGAAGTAGTG
AGGAGGCTTTTTTGGAGGCCTAGGCTTTTGCAAAAAGCTTGATTCTTCTGACACAACAGT
CTCGAACTTAAGGCTAGAGCCACCATGATTGAACAAGATGGATTGCACGCAGGTTCTCCG
GCCGCTTGGGTGGAGAGGCTATTCGGCTATGACTGGGCACAACAGACAATCGGCTGCTCT
GATGCCGCCGTGTTCCGGCTGTCAGCGCAGGGGCGCCCGGTTCTTTTTGTCAAGACCGAC
CTGTCCGGTGCCCTGAATGAACTGCAGGACGAGGCAGCGCGGCTATCGTGGCTGGCCACG
ACGGGCGTTCCTTGCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGGACTGGCTG
CTATTGGGCGAAGTGCCGGGGCAGGATCTCCTGTCATCTCACCTTGCTCCTGCCGAGAAA
GTATCCATCATGGCTGATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCA
TTCGACCACCAAGCGAAACATCGCATCGAGCGAGCACGTACTCGGATGGAAGCCGGTCTT
GTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCGCC
AGGCTCAAGGCGCGCATGCCCGACGGCGAGGATCTCGTCGTGACCCATGGCGATGCCTGC
TTGCCGAATATCATGGTGGAAAATGGCCGCTTTTCTGGATTCATCGACTGTGGCCGGCTG
GGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTT
GGCGGCGAATGGGCTGACCGCTTCCTCGTGCTTTACGGTATCGCCGCTCCCGATTCGCAG
CGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGCGGGACTCTGGGGTTCGAAA
TGACCGACCAAGCGACGCCCAACCTGCCATCACGATGGCCGCAATAAAATATCTTTATTT
TCATTACATCTGTGTGTTGGTTTTTTGTGTGAATCGATAGCGATAAGGATCCGCGTATGG
TGCACTCTCAGTACAATCTGCTCTGATGCCGCATAGTTAAGCCAGCCCCGACACCCGCCA
ACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGACAAGCT
GTGACCGTCTCCGGGAGCTGCATGTGTCAGAGGTTTTCACCGTCATCACCGAAACGCGCG
AGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAATGGTT
TCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTT
TTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAA
TAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTT
TTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGAT

GCTGAAGATCAGTTGGGTGCACGAGTGGGTTACATCGAACTGGATCTCAACAGCGGTAAG
ATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGTTCTG
CTATGTGGCGCGGTATTATCCCGTATTGACGCCGGGCAAGAGCAACTCGGTCGCCGCATA
CACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTACGGAT
GGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAACACTGCGGCC
AACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAACATG
GGGGATCATGTAACTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACCAAAC
GACGAGCGTGACACCACGATGCCTGTAGCAATGGCAACAACGTTGCGCAAACTATTAACT
GGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGATAAA
GTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAAATCT
GGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAAGCCC
TCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAATAGA
CAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGTTTAC
TCATATATACTTTAGATTGATTTAAAACTTCATTTTTAATTTAAAAGGATCTAGGTGAAG
ATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTGAGCG
TCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGTAATC
TGCTGCTTGCAAACAAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCAAGAG
CTACCAACTCTTTTTCCGAAGGTAACTGGCTTCAGCAGAGCGCAGATACCAAATACTGTC
CTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTACATAC
CTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCTTACC
GGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCGGGCTGAACGGGGGGT
TCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACAGCGT
GAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGTAAGC
GGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTATCTT
TATAGTCCTGTCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTCGTCA
GGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGCCTTT
TGCTGGCCTTTTGCTCACATGGCTCGACAGATCT

## MLV construct CZCG (SEQ ID No 55)

GTTACCTTCTGCTCTGCAGAATGGCCAACCTTTAACGTCGGATGGCCGCGAGACGGCACC
TTTAACCGAGACCTCATCACCCAGGTTAAGATCAAGGTCTTTTCACCTGGCCCGCATGGA
CACCCAGACCAGGTCCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGACCCCCCTCCC
TGGGTCAAGCCCTTTGTACACCCTAAGCCTCCGCCTCCTCTTCCTCCATCCGCCCCGTCT
CTCCCCCTTGAACCTCCTCGTTCGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACT
CCTTCTCTAGGCGCCGGAATTCGTTAACTCGAGAGGCCTGCCACCATGGGGACTGCTCCA
AAGAAGAAGCGTAAGGTAGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACC
CAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCC
CGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGG
TTTCCGGCACCAGAAGCGGTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGAT
ACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGATGCGCCCATCTACACCAAC
GTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGT
TACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATT
TTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTCGGTTACGGC
CAGGACAGTCGTTTGCCGTCTGAATTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAAC
CGCCTCGCGGTGATGGTGCTGCGTTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATG
TGGCGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCTGCATAAACCGACTACACAAATC
AGCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTCAGCCGCGCTGTACTGGAGGCT
GAAGTTCAGATGTGCGGCGAGTTGCGTGACTACCTACGGGTAACAGTTTCTTTATGGCAG
GGTGAAACGCAGGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGAGCGT
GGTGGTTATGCCGATCGCGTCACACTACGTCTGAACGTCGAAAACCCGAAACTGTGGAGC
GCCGAAATCCCGAATCTCTATCGTGCGGTGGTTGAACTGCACACCGCCGACGGCACGCTG
ATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCGGATTGAAAATGGTCTGCTG
CTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAACCGTCACGAGCATCATCCTCTG
CATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAAC
AACTTTAACGCCGTGCGCTGTTCGCATTATCCGAACCATCCGCTGTGGTACACGCTGTGC
GACCGCTACGGCCTGTATGTGGTGGATGAAGCCAATATTGAAACCCACGGCATGGTGCCA
ATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGATGAGCGAACGCGTAACGCGA
ATGGTGCAGCGCGATCGTAATCACCCGAGTGTGATCATCTGGTCGCTGGGGAATGAATCA
GGCCACGGCGCTAATCACGACGCGCTGTATCGCTGGATCAAATCTGTCGATCCTTCCCGC

CCGGTGCAGTATGAAGGCGGCGGAGCCGACACCACGGCCACCGATATTATTTGCCCGATG
TACGCGCGCGTGGATGAAGACCAGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAAAA
TGGCTTTCGCTACCTGGAGAGACGCGCCCGCTGATCCTTTGCGAATACGCCCACGCGATG
GGTAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTA
CAGGGCGGCTTCGTCTGGGACTGGGTGGATCAGTCGCTGATTAAATATGATGAAAACGGC
AACCCGTGGTCGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTTCTGT
ATGAACGGTCTGGTCTTTGCCGACCGCACGCCGCATCCAGCGCTGACGGAAGCAAAACAC
CAGCAGCAGTTTTTCCAGTTCCGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATAC
CTGTTCCGTCATAGCGATAACGAGCTCCTGCACTGGATGGTGGCGCTGGATGGTAAGCCG
CTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAGGTAAACAGTTGATTGAACTG
CCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGGCTCACAGTACGCGTAGTGCAA
CCGAACGCGACCGCATGGTCAGAAGCCGGGCACATCAGCGCCTGGCAGCAGTGGCGTCTG
GCGGAAAACCTCAGTGTGACGCTCCCCGCCGCGTCCCACGCCATCCCGCATCTGACCACC
AGCGAAATGGATTTTTGCATCGAGCTGGGTAATAAGCGTTGGCAATTTAACCGCCAGTCA
GGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACTGCTGACGCCGCTGCGCGAT
CAGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGCATTGAC
CCTAACGCCTGGGTCGAACGCTGGAAGGCGGCGGGCCATTACCAGGCCGAAGCAGCGTTG
TTGCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATTACGACCGCTCACGCGTGG
CAGCATCAGGGGAAAACCTTATTTATCAGCCGGAAAACCTACCGGATTGATGGTAGTGGT
CAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCGGCGCGGATT
GGCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCG
CAAGAAAACTATCCCGACCGCCTTACTGCCGCCTGTTTTGACCGCTGGGATCTGCCATTG
TCAGACATGTATACCCCGTACGTCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGCGC
GAATTGAATTATGGCCCACACCAGTGGCGCGGCGACTTCCAGTTCAACATCAGCCGCTAC
AGTCAACAGCAACTGATGGAAACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACA
TGGCTGAATATCGACGGTTTCCATATGGGGATTGGTGGCGACGACTCCTGGAGCCCGTCA
GTATCGGCGGAATTCCAGCTGAGCGCCGGTCGCTACCATTACCAGTTGGTCTGGTGTCAA
AAATAATAATAACCGGGCAGGGGGGATCCGCAGATCCGGCTGTGGAATGTGTGTCAGTTA
GGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCCTGCAGGA
GTGGGGAGGCACGATGGCCGCTTTGGTCGAGGCGGATCCGGCCATTAGCCATATTATTCA
TTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCCATATC
ATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCATGTTGACATTGATTAT
TGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGT
TCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCC
CATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGAC
GTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATA
TGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC
AGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTA
TTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAC
GGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC
AACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGC
ATGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGA
GACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCG
GCCCCAAGCTTGTTGGGATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTT
CACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAG
CGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTG
CACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGT
GCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCAT
GCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC
CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCAT
CGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCA
CAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCG
CCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCAT
CGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAG
CAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCGACTCTAGATCATAATC
AGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTG
AACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACATCGATAAAATAAAAGATTTT
ATTTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACCTGTAGGTTTGGCAAGCTAGC
TTAAGTAACGCCATTTTGCAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGTTCAG

ATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAG
CAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAACAG
GATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGTCCCCAGAT
GCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGGA
CCTGAAATGACCCTGTGCCTTATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTC
GCGCGCTTCTGCTCCCCGAGCTCAATAAAAGAGCCCACAACCCCTCACTCGGGGCGCCAG
TCCTCCGATTGACTGAGTCGCCCGGGTACCCGTGTATCCAATAAACCCTCTTGCAGTTGC
ATCCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGATTGACTACCCGT
CAGCGGGGGTCTTTCATTTGGGGGCTCGTCCGGGATCGGGAGACCCCTGCCCAGGGACCA
CCGACCCACCACCGGGAGGTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTA
GTGTCTATGACTGATTTTATGCGCCTGCGTCGGTACTAGTTAGCTAACTAGCTCTGTATC
TGGCGGACCCGTGGTGGAACTGACGAGTTCGGAACACCCGGCCGCAACCCTGGGAGAGGA
ATTCTCATGTTTGACAGCTTATCATCGATAAGCTTTTTGCAAAAGCCTAGGCCTCCAAAA
AAGCCTCCTCACTACTTCTGGAATAGCTCAGAGGCCGAGGCGGCCTCGGCCTCTGCATAA
ATAAAAAAAATTAGTCAGCCATGGGGCGGAGAATGGGCGGAACTGGGCGGAGTTAGGGGC
GGGATGGGCGGAGTTAGGGGCGGGACTATGGTTGCTGACTAATTGAGATGCATGCTTTGC
ATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCTGGTTGCTGACTAATTGAGA
TGCATGCTTTGCATACTTCTGCCTGCTGGGGAGCCTGGGGACTTTCCACACCCTAACTGA
CACACATTCCACAGCCGGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCC
ACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGGGCTCGC
CACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGG
GGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGC
CTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGACCG
ATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATC
GTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCG
CTCTGGGTCATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCG
CTTGCGGTATTCGGAATCTTGCACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACC
AAACGTTTCGGCGAGAAGCAGGCCATTATCGCCGGCATGGCGGCCGACGCGCTGGGCTAC
GTCTTGCTGGCGTTCGCGACGCGAGGCTGGATGGCCTTCCCCATTATGATTCTTCTCGCT
TCCGGCGGCATCGGGATGCCCGCGTTGCAGGCCATGCTGTCCAGGCAGGTAGATGACGAC
CATCAGGGACAGCTTCAAGGATCGCTCGCGGCTCTTACCAGCCTAACTTCGATCACTGGA
CCGCTGATCGTCACGGCGATTTATGCCGCCTCGGCGAGCACATGGAACGGGTTGGCATGG
ATTGTAGGCGCCGCCCTATACCTTGTCTGCCTCCCCGCGTTGCGTCGCGGTGCATGGAGC
CGGGCCACCTCGACCTGAATGGAAGCCGGCGGCACCTCGCTAACGGATTCACCACTCCAA
GAATTGGAGCCAATCAATTCTTGCGGAGAACTGTGAATGCGCAAACCAACCCTTGGCAGA
ACATATCCATCGCGTCCGCCATCTCCAGCAGCCGCACGCGGCGCATCTCGGGCAGCGTTG
GGTCCTGGCCACGGGTGCGCATGATCGTGCTCCTGTCGTTGAGGACCCGGCTAGGCTGGC
GGGGTTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGAGCGAACGTGAAGCGACT
GCTGCTGCAAAACGTCTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTTCCGTGTTT
CGTAAAGTCTGGAAACGCGGAAGTCAGCGCCCTGCACCATTATGTTCCGGATCTGCATCG
CAGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATTAACGAAGCGCTGGCATT
GACCCTGAGTGATTTTTCTCTGGTCCCGCCGCATCCATACCGCCAGTTGTTTACCCTCAC
AACGTTCCAGTAACCGGGCATGTTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTC
GTTTCATCGGTATCATTACCCCCATGAACAGAAATTCCCCCTTACACGGAGGCATCAAGT
GACCAAACAGGAAAAAACCGCCCTTAACATGGCCCGCTTTATCAGAAGCCAGACATTAAC
GCTTCTGGAGAAACTCAACGAGCTGGACGCGGATGAACAGGCAGACATCTGTGAATCGCT
TCACGACCACGCTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGA
AAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGG
GAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCAT
GACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAG
ATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAA
TACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGG
CTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGG
GATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAG
GCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGA
CGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCT
GGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCC
TTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCG
GTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGC
TGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCA

CTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAG
TTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCT
CTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACC
ACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGA
TCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCA
CGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAAT
TAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTAC
CAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTT
GCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGT
GCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAG
CCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCT
ATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTT
GTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGC
TCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTT
AGCTCCTTCGGTCCTCCGATCGTCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCT
GTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGC
TCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTC
ATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCC
AGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGC
GTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA
CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGT
TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTT
CCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACA
TTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGAC
GGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGAT
GCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGG
CTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATCGACGCTCTCCCT
TATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCG
CCGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTG
CCACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCC
CATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGG
CCACGATGCGTCCGGCGTAGAGGATCTGGCTAGCGATGACCCTGCTGATTGGTTCGCTGA
CCATTTCCGGGGTGCGGAACGGCGTTACCAGAAACTCAGAAGGTTCGTCCAACCAAACCG
ACTCTGACGGCAGTTTACGAGAGAGATGATAGGGTCTGCTTCAGTAAGCCAGATGCTACA
CAATTAGGCTTGTACATATTGTCGTTAGAACGCGGCTACAATTAATACATAACCTTATGT
ATCATACACATACGATTTAGGTGACACTATAGAATACAAGCTGGAAGATCTTCCAGCTTG
GGCTGCAGGTCGACTCTAGAGTCCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGA
CCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA
ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCA
GTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGG
CCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATC
TACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGT
GGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGT
TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTG
ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG
AACCGCGCCAGTCTTCCGATAGACTGCGTCGCCCGGGTACCCGTATTCCCAATAAAGCCT
CTTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGA
TTGACTACCCACGACGGGGGTCTTTCATTTGGGGGCTCGTCCGGGATTTGGAGACCCCTG
CCCAGGGACCACCGACCCACCACCGGGAGGTAAGCTGGCCAGCAACTTATCTGTGTCTGT
CCGATTGTCTAGTGTCTATGTTTGATGTTATGCGCCTGCGTCTGTACTAGTTAGCTAACT
AGCTCTGTATCTGGCGGACCCGTGGTGGAACTGACGAGTTCTGAACACCCGGCCGCAACC
CTGGGAGACGTCCCAGGGACTTTGGGGGCCGTTTTTGTGGCCCGACCTGAGGAAGGGAGT
CGATGTGGAATCCGACCCCGTCAGGATATGTGGTTCTGGTAGGAGACGAGAACCTAAAAC
AGTTCCCGCCTCCGTCTGAATTTTTGCTTTCGGTTTGGAACCGAAGCCGCGCGTCTTGTC
TGCTGCAGCGCTGCAGCATCGTTCTGTGTTGTCTCTGTCTGACTGTGTTTCTGTATTTGT
CTGAAAATTAGGGCCAGACTGTTACCACTCCCTTAAGTTTGACCTTAGGTCACTGGAAAG
ATGTCGAGCGGATCGCTCACAACCAGTCGGTAGATGTCAAGAAGAGACGTTGG

**PTRAP2 (SEQ ID No 56)**

GTTACCTTCTGCTCTGCAGAATGGCCAACCTTTAACGTCGGATGGCCGCGAGACGGCACC
TTTAACCGAGACCTCATCACCCAGGTTAAGATCAAGGTCTTTTCACCTGGCCCGCATGGA
CACCCAGACCAGGTCCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGACCCCCCTCCC
TGGGTCAAGCCCTTTGTACACCCTAAGCCTCCGCCTCCTCTTCCTCCATCCGCCCCGTCT
CTCCCCCTTGAACCTCCTCGTTCGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACT
CCTTCTCTAGGCGCCGGAATTCGTTAACTCGAGAGGCCTGCCACCATGGGGACTGCTCCA
AAGAAGAAGCGTAAGGTAGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACC
CAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCC
CGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGCGCTTTGCCTGG
TTTCCGGCACCAGAAGCGGTGCCGGAAAGCTGGCTGGAGTGCGATCTTCCTGAGGCCGAT
ACTGTCGTCGTCCCCTCAAACTGGCAGATGCACGGTTACGATGCGCCCATCTACACCAAC
GTAACCTATCCCATTACGGTCAATCCGCCGTTTGTTCCCACGGAGAATCCGACGGGTTGT
TACTCGCTCACATTTAATGTTGATGAAAGCTGGCTACAGGAAGGCCAGACGCGAATTATT
TTTGATGGCGTTAACTCGGCGTTTCATCTGTGGTGCAACGGGCGCTGGGTCGGTTACGGC
CAGGACAGTCGTTTGCCGTCTGAATTTGACCTGAGCGCATTTTTACGCGCCGGAGAAAAC
CGCCTCGCGGTGATGGTGCTGCGTTGGAGTGACGGCAGTTATCTGGAAGATCAGGATATG
TGGCGGATGAGCGGCATTTTCCGTGACGTCTCGTTGCTGCATAAACCGACTACACAAATC
AGCGATTTCCATGTTGCCACTCGCTTTAATGATGATTTCAGCCGCGCTGTACTGGAGGCT
GAAGTTCAGATGTGCGGCGAGTTGCGTGACTACCTACGGGTAACAGTTTCTTTATGGCAG
GGTGAAACGCAGGTCGCCAGCGGCACCGCGCCTTTCGGCGGTGAAATTATCGATGAGCGT
GGTGGTTATGCCGATCGCGTCACACTACGTCTGAACGTCGAAAACCCGAAACTGTGGAGC
GCCGAAATCCCGAATCTCTATCGTGCGGTGGTTGAACTGCACACCGCCGACGGCACGCTG
ATTGAAGCAGAAGCCTGCGATGTCGGTTTCCGCGAGGTGCGGATTGAAAATGGTCTGCTG
CTGCTGAACGGCAAGCCGTTGCTGATTCGAGGCGTTAACCGTCACGAGCATCATCCTCTG
CATGGTCAGGTCATGGATGAGCAGACGATGGTGCAGGATATCCTGCTGATGAAGCAGAAC
AACTTTAACGCCGTGCGCTGTTCGCATTATCCGAACCATCCGCTGTGGTACACGCTGTGC
GACCGCTACGGCCTGTATGTGGTGGATGAAGCCAATATTGAAACCCACGGCATGGTGCCA
ATGAATCGTCTGACCGATGATCCGCGCTGGCTACCGGCGATGAGCGAACGCGTAACGCGA
ATGGTGCAGCGCGATCGTAATCACCCGAGTGTGATCATCTGGTCGCTGGGGAATGAATCA
GGCCACGGCGCTAATCACGACGCGCTGTATCGCTGGATCAAATCTGTCGATCCTTCCCGC
CCGGTGCAGTATGAAGGCGGCGGAGCCGACACCACGGCCACCGATATTATTTGCCCGATG
TACGCGCGCGTGGATGAAGACCAGCCCTTCCCGGCTGTGCCGAAATGGTCCATCAAAAAA
TGGCTTTCGCTACCTGGAGAGACGCGCCCGCTGATCCTTTGCGAATACGCCCACGCGATG
GGTAACAGTCTTGGCGGTTTCGCTAAATACTGGCAGGCGTTTCGTCAGTATCCCCGTTTA
CAGGGCGGCTTCGTCTGGGACTGGGTGGATCAGTCGCTGATTAAATATGATGAAAACGGC
AACCCGTGGTCGGCTTACGGCGGTGATTTTGGCGATACGCCGAACGATCGCCAGTTCTGT
ATGAACGGTCTGGTCTTTGCCGACCGCACGCCGCATCCAGCGCTGACGGAAGCAAAACAC
CAGCAGCAGTTTTTCCAGTTCCGTTTATCCGGGCAAACCATCGAAGTGACCAGCGAATAC
CTGTTCCGTCATAGCGATAACGAGCTCCTGCACTGGATGGTGGCGCTGGATGGTAAGCCG
CTGGCAAGCGGTGAAGTGCCTCTGGATGTCGCTCCACAAGGTAAACAGTTGATTGAACTG
CCTGAACTACCGCAGCCGGAGAGCGCCGGGCAACTCTGGCTCACAGTACGCGTAGTGCAA
CCGAACGCGACCGCATGGTCAGAAGCCGGGCACATCAGCGCCTGGCAGCAGTGGCGTCTG
GCGGAAAACCTCAGTGTGACGCTCCCCGCCGCGTCCCACGCCATCCCGCATCTGACCACC
AGCGAAATGGATTTTTGCATCGAGCTGGGTAATAAGCGTTGGCAATTTAACCGCCAGTCA
GGCTTTCTTTCACAGATGTGGATTGGCGATAAAAAACAACTGCTGACGCCGCTGCGCGAT
CAGTTCACCCGTGCACCGCTGGATAACGACATTGGCGTAAGTGAAGCGACCCGCATTGAC
CCTAACGCCTGGGTCGAACGCTGGAAGGCGGCGGGCCATTACCAGGCCGAAGCAGCGTTG
TTGCAGTGCACGGCAGATACACTTGCTGATGCGGTGCTGATTACGACCGCTCACGCGTGG
CAGCATCAGGGGAAAACCTTATTTATCAGCCGGAAAACCTACCGGATTGATGGTAGTGGT
CAAATGGCGATTACCGTTGATGTTGAAGTGGCGAGCGATACACCGCATCCGGCGCGGATT
GGCCTGAACTGCCAGCTGGCGCAGGTAGCAGAGCGGGTAAACTGGCTCGGATTAGGGCCG
CAAGAAAACTATCCCGACCGCCTTACTGCCGCCTGTTTTGACCGCTGGGATCTGCCATTG
TCAGACATGTATACCCCGTACGTCTTCCCGAGCGAAAACGGTCTGCGCTGCGGGACGCGC
GAATTGAATTATGGCCCACACCAGTGGCGCGGCGACTTCCAGTTCAACATCAGCCGCTAC
AGTCAACAGCAACTGATGGAAACCAGCCATCGCCATCTGCTGCACGCGGAAGAAGGCACA
TGGCTGAATATCGACGGTTTCCATATGGGGATTGGTGGCGACGACTCCTGGAGCCCGTCA
GTATCGGCGGAATTCCAGCTGAGCGCCGGTCGCTACCATTACCAGTTGGTCTGGTGTCAA
AAATAATAATAACCGGGCAGGGGGGATCCGCAGATCCGGCTGTGGAATGTGTGTCAGTTA
GGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAAGCATGCCTGCAGGA

GTGGGGAGGCACGATGGCCGCTTTGGTCGAGGCGGATCCGGCCATTAGCCATATTATTCA
TTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCCATATC
ATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGCCATGTTGACATTGATTAT
TGACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGT
TCCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCC
CATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGAC
GTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATA
TGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC
AGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTA
TTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCAC
GGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC
AACGGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGC
ATGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGA
GACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCG
GCCCCAAGCTTGTTGGGATCCACCGGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTT
CACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAG
CGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTG
CACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACCACCCTGACCTACGGCGT
GCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGACTTCTTCAAGTCCGCCAT
GCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGACGACGGCAACTACAAGAC
CCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGCATCGAGCTGAAGGGCAT
CGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCA
CAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAGGTGAACTTCAAGATCCG
CCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTACCAGCAGAACACCCCCAT
CGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGCACCCAGTCCGCCCTGAG
CAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAGTTCGTGACCGCCGCCGG
GATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCGCGACTCTAGATCATAATC
AGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCTCCCCCTG
AACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACATCGATAAAATAAAAGATTTT
ATTTAGTCTCCAGAAAAAGGGGGGAATGAAAGACCCCACCTGTAGGTTTGGCAAGCTAGC
ATAACTTCGTATAATGTATGCTATACGAAGTTATTCTAGAGAACCATCAGATGTTTCCAG
GGTGCCCCAAGGACCTGAAATGACCCTGTGCCTTATTTGAACTAACCAATCAGTTCGCTT
CTCGCTTCTGTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAAGAGCCCACAACCCCTCA
CTCGGGGCGCCAGTCCTCCGATTGACTGAGTCGCCCGGGTACCCGTGTATCCAATAAACC
CTCTTGCAGTTGCATCCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTG
ATTGACTACCCGTCAGCGGGGGTCTTTCATTTGGGGGCTCGTCCGGGATCGGGAGACCCC
TGCCCAGGGACCACCGACCCACCACCGGGAGGTAAGCTGGCTGCCTCGCGCGTTTCGGTG
ATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAG
CGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGG
GCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGC
ATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGT
AAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTC
GGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCAC
AGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAA
CCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCA
CAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGC
GTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATA
CCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTA
TCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCA
GCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGA
CTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGG
TGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGG
TATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGG
CAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAG
AAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAA
CGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGAT
CCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTC
TGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTC
ATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATC
TGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGC

AATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTC
CATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTT
GCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGC
TTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAA
AAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTT
ATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATG
CTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACC
GAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAA
AGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTT
GAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTT
CACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAG
GGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTA
TCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAAT
AGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTAT
CATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATTCAT
ACCAGATCACCGAAAACTGTCCTCCAAATGTGTCCCCCTCACACTCCCAAATTCGCGGGC
TTCTGCCTCTTAGACCACTCTACCCTATTCCCCACACTCACCGGAGCCAAAGCCGCGGCC
CTTCCGTTTCTTTGCTTTTGAAAGACCCCACCCGTAGGTGGCAAGCTAGCGATGACCCTG
CTGATTGGTTCGCTGACCATTTCCGGGGTGCGGAACGGCGTTACCAGAAACTCAGAAGGT
TCGTCCAACCAAACCGACTCTGACGGCAGTTTACGAGAGAGATGATAGGGTCTGCTTCAG
TAAGCCAGATGCTACACAATTAGGCTTGTACATATTGTCGTTAGAACGCGGCTACAATTA
ATACATAACCTTATGTATCATACACATACGATTTAGGTGACACTATAGAATACAAGCTGG
AAGATCTTCCAGCTTGGGCTGCAGGTCGACTCTAGAGTCCGTTACATAACTTACGGTAAA
TGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGT
TCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTA
AACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGT
CAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCC
TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCA
GTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCAT
TGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA
CAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG
CAGAGCTCGTTTAGTGAACCGCGCCAGTCTTCCGATAGACTGCGTCGCCCGGGTACCCGT
ATTCCCAATAAAGCCTCTTGCTGTTTGCATCCGAATCGTGGTCTCGCTGTTCCTTGGGAG
GGTCTCCTCTGAGTGATTGACTACCCACGACGGGGGTCTTTCATTTGGGGGCTCGTCCGG
GATTTGGAGACCCCTGCCCAGGGACCACCGACCCACCACCGGGAGGTAAGCTGGCCAGCA
ACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGTTTGATGTTATGCGCCTGCGTCTG
TACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCGTGGTGGAACTGACGAGTTCTGA
ACACCCGGCCGCAACCCTGGGAGACGTCCCAGGGACTTTGGGGGCCGTTTTTGTGGCCCG
ACCTGAGGAAGGGAGTCGATGTGGAATCCGACCCCGTCAGGATATGTGGTTCTGGTAGGA
GACGAGAACCTAAAACAGTTCCCGCCTCCGTCTGAATTTTTGCTTTCGGTTTGGAACCGA
AGCCGCGCGTCTTGTCTGCTGCAGCGCTGCAGCATCGTTCTGTGTTGTCTCTGTCTGACT
GTGTTTCTGTATTTGTCTGAAAATTAGGGCCAGACTGTTACCACTCCCTTAAGTTTGACC
TTAGGTCACTGGAAAGATGTCGAGCGGATCGCTCACAACCAGTCGGTAGATGTCAAGAAG
AGACGTTGG

**PCGCLNGFR (SEQ ID No 57)**

GTTACCTTCTGCTCTGCAGAATGGCCAACCTTTAACGTCGGATGGCCGCGAGACGGCACC
TTTAACCGAGACCTCATCACCCAGGTTAAGATCAAGGTCTTTTCACCTGGCCCGCATGGA
CACCCAGACCAGGTCCCCTACATCGTGACCTGGGAAGCCTTGGCTTTTGACCCCCCTCCC
TGGGTCAAGCCCTTTGTACACCCTAAGCCTCCGCCTCCTCTTCCTCCATCCGCCCCGTCT
CTCCCCCTTGAACCTCCTCGTTCGACCCCGCCTCGATCCTCCCTTTATCCAGCCCTCACT
CCTTCTCTAGGCGCCGGAATTCGTTAACTCGAGGATCCACCGGTCGCCACCATGGTGAGC
AAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA
AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTG
ACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACC
ACCCTGACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGAAGCAGCACGAC
TTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCTTCTTCAAGGAC
GACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCCTGGTGAACCGC
ATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACAAGCTGGAG

TACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGAACGGCATCAAG
GTGAACTTCAAGATCCGCCACAACATCGAGGACGGCAGCGTGCAGCTCGCCGACCACTAC
CAGCAGAACACCCCCATCGGCGACGGCCCCGTGCTGCTGCCCGACAACCACTACCTGAGC
ACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGGTCCTGCTGGAG
TTCGTGACCGCCGCCGGGATCACTCTCGGCATGGACGAGCTGTACAAGTAAAGCGGCCCT
AGGGGTCTTTCCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGGAAGCA
GTTCCTCTGGAAGCTTCTTGAAGACAAACAACGTCTGTAGCGACCCTTTGCAGGCAGCGG
AACCCCCCACCTGGCGACAGGTGCCTCTGCGGCCAAAAGCCACCGAGTTGGTTCAGCTGC
TGCCTGAGGCTGGACGACCTCGCGGAGTTCTACCGGCAGTGCAAATCCGTCGGCATCCAG
GAAACCAGCAGCGGCTATCCGCGCATCCATGCCCCCGAACTGCAGGAGTGGGGAGGCACG
ATGGCCGCTTTGGTCGAGGCGGATCCGGCCATTAGCCATATTATTCATTGGTTATATAGC
ATAAATCAATATTGGCTATTGGCCATTGCATACGTTGTATCCATATCATAATATGTACAT
TTATATTGGCTCATGTCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTTATTA
ATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATA
ACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAAT
AATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGA
GTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCC
CCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT
ATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGAT
GCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG
TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCC
AAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCATGTACGGTGGGA
GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAGACGCCATCCACG
CTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGGCCCCAAGCTTAC
CATGGGGGCAGGTGCCACCGGCCGCGCCATGGACGGGCCGCGCCTGCTGCTGTTGCTGCT
TCTGGGGGTGTCCCTTGGAGGTGCCAAGGAGGCATGCCCCACAGGCCTGTACACACACAG
CGGTGAGTGCTGCAAAGCCTGCAACCTGGGCGAGGGTGTGGCCCAGCCTTGTGGAGCCAA
CCAGACCGTGTGTGAGCCCTGCCTGGACAGCGTGACGTTCTCCGACGTGGTGAGCGCGAC
CGAGCCGTGCAAGCCGTGCACCGAGTGCGTGGGGCTCCAGAGCATGTCGGCGCCGTGCGT
GGAGGCCGACGACGCCGTGTGCCGCTGCGCCTACGGCTACTACCAGGATGAGACGACTGG
GCGCTGCGAGGCGTGCCGCGTGTGCGAGGCGGGCTCGGGCCTCGTGTTCTCCTGCCAGGA
CAAGCAGAACACCGTGTGCGAGGAGTGCCCCGACGGCACGTATTCCGACGAGGCCAACCA
CGTGGACCCGTGCCTGCCCTGCACCGTGTGCGAGGACACCGAGCGCCAGCTCCGCGAGTG
CACACGCTGGGCCGACGCCGAGTGCGAGGAGATCCCTGGCCGTTGGATTACACGGTCCAC
ACCCCCAGAGGGCTCGGACAGCACAGCCCCCAGCACCCAGGAGCCTGAGGCACCTCCAGA
ACAAGACCTCATAGCCAGCACGGTGGCAGGTGTGGTGACCACAGTGATGGGCAGCTCCCA
GCCCGTGGTGACCCGAGGCACCACCGACAACCTCATCCCTGTCTATTGCTCCATCCTGGC
TGCTGTGGTTGTGGGCCTTGTGGCCTACATAGCCTTCAAGAGGTGGAACAGCTGCTGAGT
CGACTCTAGAGGATCCCCAACATCGATAAAATAAAAGATTTTATTTAGTCTCCAGAAAAA
GGGGGGAATGAAAGACCCCACCTGTAGGTTTGGCAAGCTAGCTTAAGTAACGCCATTTTG
CAAGGCATGGAAAAATACATAACTGAGAATAGAGAAGTTCAGATCAAGGTCAGGAACAGA
TGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTC
AGGGCCAAGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGTGGTAAGCAG
TTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGTCCCCAGATGCGGTCCAGCCCTCAGCA
GTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCCCAAGGACCTGAAATGACCCTGTGC
CTTATTTGAACTAACCAATCAGTTCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCCG
AGCTCAATAAAAGAGCCCACAACCCCTCACTCGGGGCGCCAGTCCTCCGATTGACTGAGT
CGCCCGGGTACCCGTGTATCCAATAAACCCTCTTGCAGTTGCATCCGACTTGTGGTCTCG
CTGTTCCTTGGGAGGGTCTCCTCTGAGTGATTGACTACCCGTCAGCGGGGGTCTTTCATT
TGGGGGCTCGTCCGGGATCGGGAGACCCCTGCCCAGGGACCACCGACCCACCACCGGGAG
GTAAGCTGGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCT
CCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGG
CGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAG
CGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCAT
ATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCC
GCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCT
CACTCAAAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATG
TGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTC
CATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGA
AACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCT

CCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTG
GCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAG
CTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTAT
CGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAAC
AGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAAC
TACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTC
GGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTT
TTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATC
TTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATG
AGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCA
ATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCA
CCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAG
ATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGAC
CCACGCTCACCGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGC
AGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCT
AGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATC
GTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGG
CGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATC
GTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAAT
TCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAG
TCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGAT
AATACCGCGCCACATAGCAGAACTTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGG
CGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCA
CCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGA
AGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATA
TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTG
CCACCTGACGTCTAAGAAACCATTATTATCATGACATTAACCTATAAAAATAGGCGTATC
ACGAGGCCCTTTCGTCTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAG
CTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAG
GGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTATGCGGCATCAGAGCAG
ATTGTACTGAGAGTGCACCATATGGACATATTGTCGTTAGAACGCGGCTACAATTAATAC
ATAACCTTATGTATCATACACATACGATTTAGGTGACACTATAGAACTCGACTCTAGAGT
CCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCAT
TGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTC
AATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGC
CAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGT
ACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTA
CCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACGGG
GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAAC
GGGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTG
TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGCGCCAGTCTTCCGATAG
ACTGCGTCGCCCGGGTACCCGTATTCCCAATAAAGCCTCTTGCTGTTTGCATCCGAATCG
TGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGATTGACTACCCACGACGGGGGTC
TTTCATTTGGGGGCTCGTCCGGGATTTGGAGACCCCTGCCCAGGGACCACCGACCCACCA
CCGGGAGGTAAGCTGGCCAGCAACTTATCTGTGTCTGTCCGATTGTCTAGTGTCTATGTT
TGATGTTATGCGCCTGCGTCTGTACTAGTTAGCTAACTAGCTCTGTATCTGGCGGACCCG
TGGTGGAACTGACGAGTTCTGAACACCCGGCCGCAACCCTGGGAGACGTCCCAGGGACTT
TGGGGGCCGTTTTTGTGGCCCGACCTGAGGAAGGGAGTCGATGTGGAATCCGACCCCGTC
AGGATATGTGGTTCTGGTAGGAGACGAGAACCTAAAACAGTTCCCGCCTCCGTCTGAATT
TTTGCTTTCGGTTTGGAACCGAAGCCGCGCGTCTTGTCTGCTGCAGCGCTGCAGCATCGT
TCTGTGTTGTCTCTGTCTGACTGTGTTTCTGTATTTGTCTGAAAATTAGGGCCAGACTGT
TACCACTCCCTTAAGTTTGACCTTAGGTCACTGGAAAGATGTCGAGCGGATCGCTCACAA
CCAGTCGGTAGATGTCAAGAAGAGACGTTGG

**PLTRloxP (SEQ ID No 58)**

GCTAGCATAACTTCGTATAATGTATGCTATACGAAGTTATTCTAGAGAACCATCAGATGT
TTCCAGGGTGCCCCAAGGACCTGAAATGACCCTGTGCCTTATTTGAACTAACCAATCAGT

TCGCTTCTCGCTTCTGTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAAGAGCCCACAAC
CCCTCACTCGGGGCGCCAGTCCTCCGATTGACTGAGTCGCCCGGGTACCCGTGTATCCAA
TAAACCCTCTTGCAGTTGCATCCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTC
TGAGTGATTGACTACCCGTCAGCGGGGGTCTTTCATTTGGGGGCTCGTCCGGGATCGGGA
GACCCCTGCCCAGGGACCACCGACCCACCACCGGGAGGTAAGCTGGCTGCCTCGCGCGTT
TCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTC
TGTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGT
GTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTA
TGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAG
ATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCT
GCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTT
ATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGC
CAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGA
GCATCACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATA
CCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTAC
CGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTG
TAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCC
CGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAG
ACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGT
ATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTG
ATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTAC
GCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCA
GTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCAC
CTAGATCCTTTTAAATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAAC
TTGGTCTGACAGTTACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATT
TCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTT
ACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTT
ATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATC
CGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAA
TAGTTTGCGCAACGTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTTTGG
TATGGCTTCATTCAGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTT
GTGCAAAAAAGCGGTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGC
AGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGT
AAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCG
GCGACCGAGTTGCTCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAAC
TTTAAAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACC
GCTGTTGAGATCCAGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTT
TACTTTCACCAGCGTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGG
AATAAGGGCGACACGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAG
CATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAA
ACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCAT
TATTATCATGACATTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGA
ATTCATACCAGATCACCGAAAACTGTCCTCCAAATGTGTCCCCCTCACACTCCCAAATTC
GCGGGCTTCTGCCTCTTAGACCACTCTACCCTATTCCCCACACTCACCGGAGCCAAAGCC
GCGGCCCTTCCGTTTCTTTGCTTTTGAAAGACCCCACCCGTAGGTGGCAA

## LTR plasmid (SEQ ID No 59)

GCTAGCTTAAGTAACGCCATTTTGCAAGGCATGGAAAAATACATAACTGAGAATAGAGAA
GTTCAGATCAAGGTCAGGAACAGATGGAACAGCTGAATATGGGCCAAACAGGATATCTGT
GGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGAACAGCTGAATATGGGCC
AAACAGGATATCTGTGGTAAGCAGTTCCTGCCCCGGCTCAGGGCCAAGAACAGATGGTCC
CCAGATGCGGTCCAGCCCTCAGCAGTTTCTAGAGAACCATCAGATGTTTCCAGGGTGCCC
CAAGGACCTGAAATGACCCTGTGCCTTATTTGAACTAACCAATCAGTTCGCTTCTCGCTT
CTGTTCGCGCGCTTCTGCTCCCCGAGCTCAATAAAAGAGCCCACAACCCCTCACTCGGGG
CGCCAGTCCTCCGATTGACTGAGTCGCCCGGGTACCCGTGTATCCAATAAACCCTCTTGC
AGTTGCATCCGACTTGTGGTCTCGCTGTTCCTTGGGAGGGTCTCCTCTGAGTGATTGACT
ACCCGTCAGCGGGGGTCTTTCATTTGGGGGCTCGTCCGGGATCGGGAGACCCCTGCCCAG
GGACCACCGACCCACCACCGGGAGGTAAGCTGGCTGCCTCGCGCGTTTCGGTGATGACGG

TGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGC
CGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGC
CATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAG
CAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGA
AAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTT
CGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCA
GGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAA
AAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAAT
CGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCC
CCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCC
GCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGT
TCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGAC
CGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCG
CCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACA
GAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGC
GCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAA
ACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAA
GGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAAC
TCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTA
AATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGT
TACCAATGCTTAATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATA
GTTGCCTGACTCCCCGTCGTGTAGATAACTACGATACGGGAGGGCTTACCATCTGGCCCC
AGTGCTGCAATGATACCGCGAGACCCACGCTCACCGGCTCCAGATTTATCAGCAATAAAC
CAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTATCCGCCTCCATCCAG
TCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCGCAAC
GTTGTTGCCATTGCTGCAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTC
AGCTCCGGTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCG
GTTAGCTCCTTCGGTCCTCCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTC
ATGGTTATGGCAGCACTGCATAATTCTCTTACTGTCATGCCATCCGTAAGATGCTTTTCT
GTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGTGTATGCGGCGACCGAGTTGC
TCTTGCCCGGCGTCAACACGGGATAATACCGCGCCACATAGCAGAACTTTAAAAGTGCTC
ATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCC
AGTTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGC
GTTTCTGGGTGAGCAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACA
CGGAAATGTTGAATACTCATACTCTTCCTTTTTCAATATTATTGAAGCATTTATCAGGGT
TATTGTCTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTT
CCGCGCACATTTCCCCGAAAAGTGCCACCTGACGTCTAAGAAACCATTATTATCATGACA
TTAACCTATAAAAATAGGCGTATCACGAGGCCCTTTCGTCTTCAAGAATTCATACCAGAT
CACCGAAAACTGTCCTCCAAATGTGTCCCCCTCACACTCCCAAATTCGCGGGCTTCTGCC
TCTTAGACCACTCTACCCTATTCCCCACACTCACCGGAGCCAAAGCCGCGGCCCTTCCGT
TTCTTTGCTTTTGAAAGACCCCACCCGTAGGTGGCAA

# CLAIMS

1. A method of modifying a producer cell which producer cell comprises integrated into its genome a provirus which provirus comprises one or more recombinase recognition sequences within or upstream of its 3' LTR, the method comprising: introducing into the cell a construct comprising a 5' recombinase recognition sequence, an LTR and a 3' recombinase recognition sequence in that order, in the presence of a recombinase which is capable of acting on the recombinase recognition site(s) such that the nucleotide sequence between the 5' and 3' recombinase recognition sequences in the construct is introduced into the provirus.

2. A method according to claim 1 wherein the construct further comprises at least one nucleotide sequence of interest (NOI) between the 5' recombinase recognition sequence and the LTR, which NOI is operably linked to a transcriptional regulatory sequence.

3. A method according to claim 1 or claim 2 wherein the construct further comprises a 5'LTR and/or a packaging signal.

4. A method according to any one of claims 1 to 3 wherein the LTR is a heterologous regulatable LTR.

5. A method according to claim 4 wherein the regulatable LTR comprises an ischaemic like response element (ILRE).

6. A method according to any one of claims 1 to 3 wherein the LTR is inactive.

7. A method according to any one of the preceding claims wherein the provirus comprises an NOI encoding a selectable marker, which NOI is flanked by recombinase recognition sites

8. A method according to any one of the precedings claims wherein the provirus comprises an internal 5' LTR upstream of the recombinase site or the 5' recombinase site

where there is more than one site.

9. A method according to any one of the preceding claims wherein the U3 region of the 5' LTR is inactive.

10. A method according to any one of the preceding claims wherein the U3 region of the 5' LTR and/or the U3 region of the second internal 5'LTR comprises a heterologous promoter.

11. A method according to any one of the preceding claims wherein the provirus comprises two recombinase recognition sites and as a preliminary step, the recombinase is expressed in a host cell such that the nucleotide sequence present between the two sites is excised.

12. A method according to any one of the preceding claims wherein the producer cell is a high titre producer cell.

13. A method according to any one of the preceding claims wherein the provirus is a lentivirus.

14. A method according to claim 13 wherein the lentivirus is HIV or EIAV.

15. A method according to any one of claims 2-14 wherein the provirus further comprises a second NOI.

16. A producer cell obtainable by the method of any one of claims 1 to 15.

17. An infectious retroviral particle obtainable from the producer cell of claim 16.

18. A derived producer cell comprising integrated into its genome a retroviral vector comprising in the 5' to 3' direction a first 5' LTR; a second NOI operably linked to a second regulatable 3' LTR; and a third 3'LTR;

wherein the third 3'LTR is positioned downstream of the second regulatable 3'LTR in the producer cell.

19. A producer cell according to claim 18 wherein the first 5' LTR comprising 5'R and 5' U5 sequences is derivable from a first vector; the second NOI operably linked to a second regulatable 3' LTR is derivable from a second vector; and the third 3'LTR is derivable from the first vector.

20. A producer cell according to claim 18 or claim 19 wherein the first vector comprises a retroviral vector wherein the retroviral vector comprises a first NOI flanked by recombinase recognition sequences.

21. A producer cell according to claim 19 or claim 20 wherein the retroviral vector further comprises an internal LTR located upstream of the first NOI and downstream of a packaging signal wherein the internal LTR comprises a heterologous U3 sequence linked to heterologous R and U5 sequences.

22. A producer cell according to any one of claims 18 to 21 wherein the third 3'LTR is transcriptionally quiescent.

23. A producer cell according to claim 22 wherein the third 3' LTR comprises a deletion in the U3 sequence.

24. A producer cell according to any one of claims 18 to 23 wherein the first NOI is a selectable marker.

25. A producer cell according to claim 19 wherein the second vector comprises a second NOI operably linked to a second regulatable 3'LTR comprising at least one recombinase recognition sequence.

26. A producer cell according to 25 wherein the second regulatable 3'LTR comprises a deletion in the U3 sequences in the 3'LTR.

27. A producer cell according to claim 25 or claim 26 wherein the second NOI comprises a coding sequence operably linked to a promoter.

28. A producer cell according to claim 27 wherein the second NOI comprises a discistronic construct.

29. A producer cell according to claim 28 wherein the discistronic construct comprises a therapeutic gene, an internal ribosomal entry site (IRES) and a reporter gene.

30. A method for producing a high titre regulatable retroviral vector, the method comprising the steps of:

(i) providing a derived producer cell comprising integrated into its genome a first vector;

(ii) introducing a second vector into the derived producer cell using a recombinase assisted method;

wherein the derived producer cell comprises a retroviral vector comprising in the 5' to 3' direction a first 5' LTR; a second NOI operably linked to a second regulatable 3' LTR; and a third 3'LTR; wherein the third 3'LTR is positioned downstream of the second regulatable 3'LTR in the derived producer cell.

31. A method according to claim 30 wherein the third 3' LTR is transcriptionally active but expression is directed away from the second regulatable 3'LTR.

32. A method for introducing a second regulatable 3'LTR into a derived producer cell wherein the method comprises a recombinase assisted method.

33. A method according to claim 32 wherein the recombinase assisted method is a Cre/lox recombinase method.

34. A process for preparing a regulated retroviral vector as defined in claim 17 comprising performing the method according to any one of claims 30 to 33 and preparing a quantity of the regulated retroviral vector.

35. A regulated retroviral vector produced by the process according to claim 34.

36. A regulated retroviral vector according to claim 35 wherein the retroviral vector is capable of transducing a target site.

37. A regulated retroviral vector according to claim 36 wherein the retroviral vector is produced in sufficient amounts to effectively transduce a target site.

38. A regulated retroviral vector according to claim 36 or claim 37 wherein the target site is a cell.

39. A cell transduced with a regulated retroviral vector according to claim 38.

40. Use of a regulated retroviral vector according to any one of claims 35 to 38 in the manufacture of a pharmaceutical composition to deliver an NOI to a target site.

41. Use of a regulated retroviral vector according to any one of claims 35 to 38 in the manufacture of a medicament for diagnostic and/or therapeutic and/or medical applications.

42. Use of a recombinase assisted mechanism to introduce a regulated 3'LTR into a derived producer cell line to produce a high titre regulated retroviral vector.

43. A derived stable producer cell capable of expressing regulated retroviral vectors according to claims 35 to 38.

44 A derived stable producer cell according to claim 43 wherein the regulated retroviral vector is a high titre regulated retroviral vector.

49. A nucleic acid molecule according to any one of claims 46 to 48 wherein the LTR is a heterologous regulatable LTR.

50. A nucleic acid molecule according to any one of claims 46 to 48 wherein the LTR is transcriptionally quiescent.

51. A method and/or a producer cell substantially as described herein and with reference to the accompanying Figures.

(12) INTERNATIONAL APP[redacted]TION PUBLISHED UNDER THE PATENT [redacted]PERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 April 2001 (12.04.2001)

PCT

(10) International Publication Number
**WO 01/25466 A1**

(51) International Patent Classification[7]: C12N 15/867, 15/90, 5/10, 7/01, A61K 48/00

(21) International Application Number: PCT/GB00/03837

(22) International Filing Date: 5 October 2000 (05.10.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
9923558.2 5 October 1999 (05.10.1999) GB

(71) Applicant *(for all designated States except US)*: **OXFORD BIOMEDICA (UK) LIMITED** [GB/GB]; Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **SLINGSBY, Jason** [GB/GB]; 91 Lacy Road, Putney, London SW15 1NR (GB). **KINGSMAN, Susan, Mary** [GB/GB]; Oxford BioMedica (UK) Limited, Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **ROHLL, Jonathan** [GB/GB]; Oxford BioMedica (UK) Limited, Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB). **SLADE, Andrew** [GB/GB]; Oxford BioMedica (UK) Limited, Medawar Centre, Robert Robinson Avenue, The Oxford Science Park, Oxford OX4 4GA (GB).

(74) Agents: **HARDING, Charles, Thomas** et al.; D Young & Co., 21 New Fetter Lane, London EC4A 1DA (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

**Published:**
— *With international search report.*
— *Before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments.*

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: PRODUCER CELL FOR THE PRODUCTION OF RETROVIRAL VECTORS

**(57) Abstract:** A method is provided for modifying a producer cell which producer cell comprises integrated into its genome a provirus which provirus comprises one or more recombinase recognition sequences within or upstream of its 3' LTR, the method comprising: introducing into the cell a construct comprising a 5' recombinase recognition sequence, an LTR and a 3' recombinase recognition sequence in that order, in the presence of a recombinase which is capable of acting on the recombinase recognition site(s) such that the nucleotide sequence between the 5' and 3' recombinase recognition sequences in the construct is introduced into the provirus.

WO 01/25466 A1

FIG. 1

MLV-based transduction using Cre/loxP system as previously described

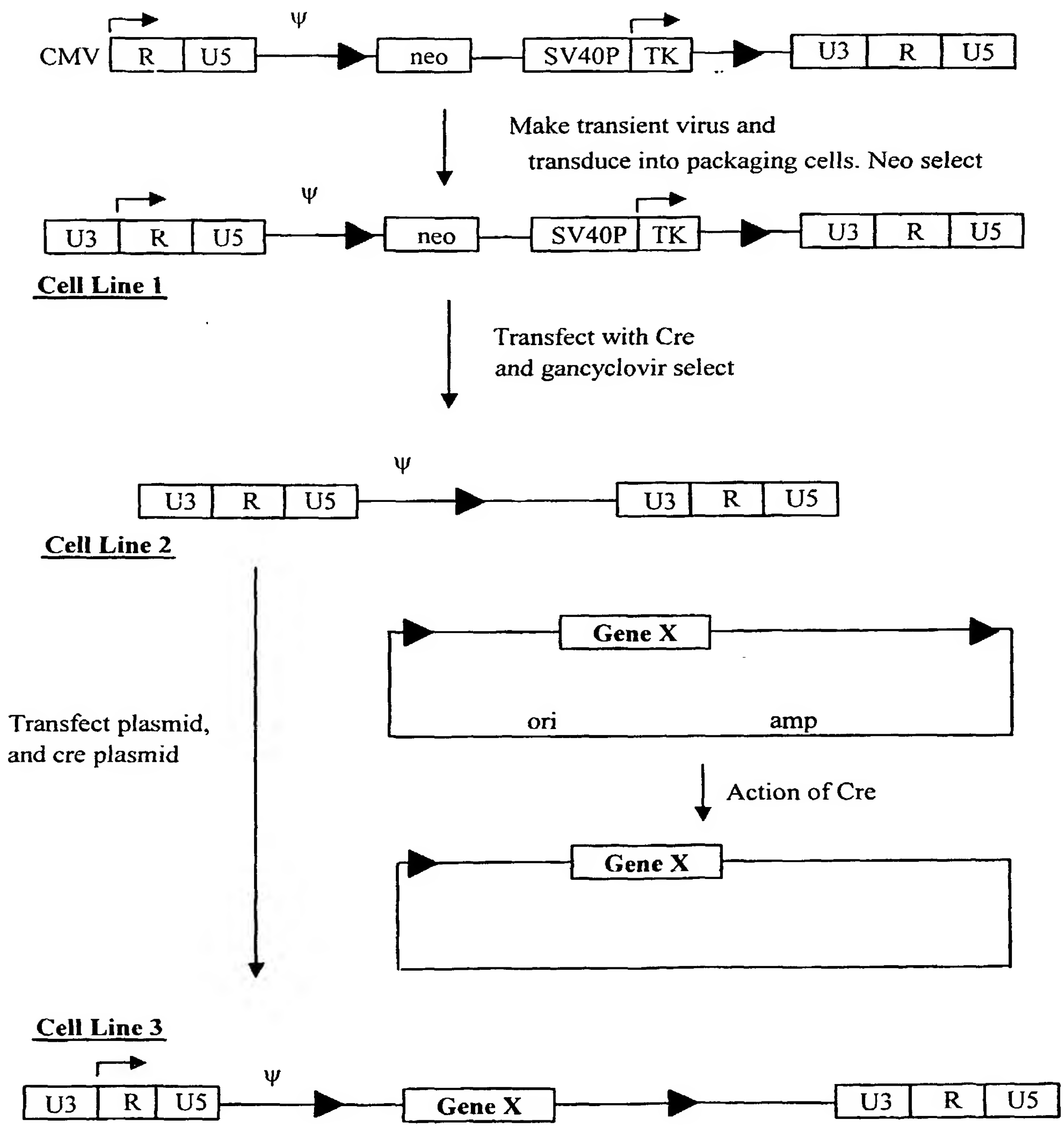
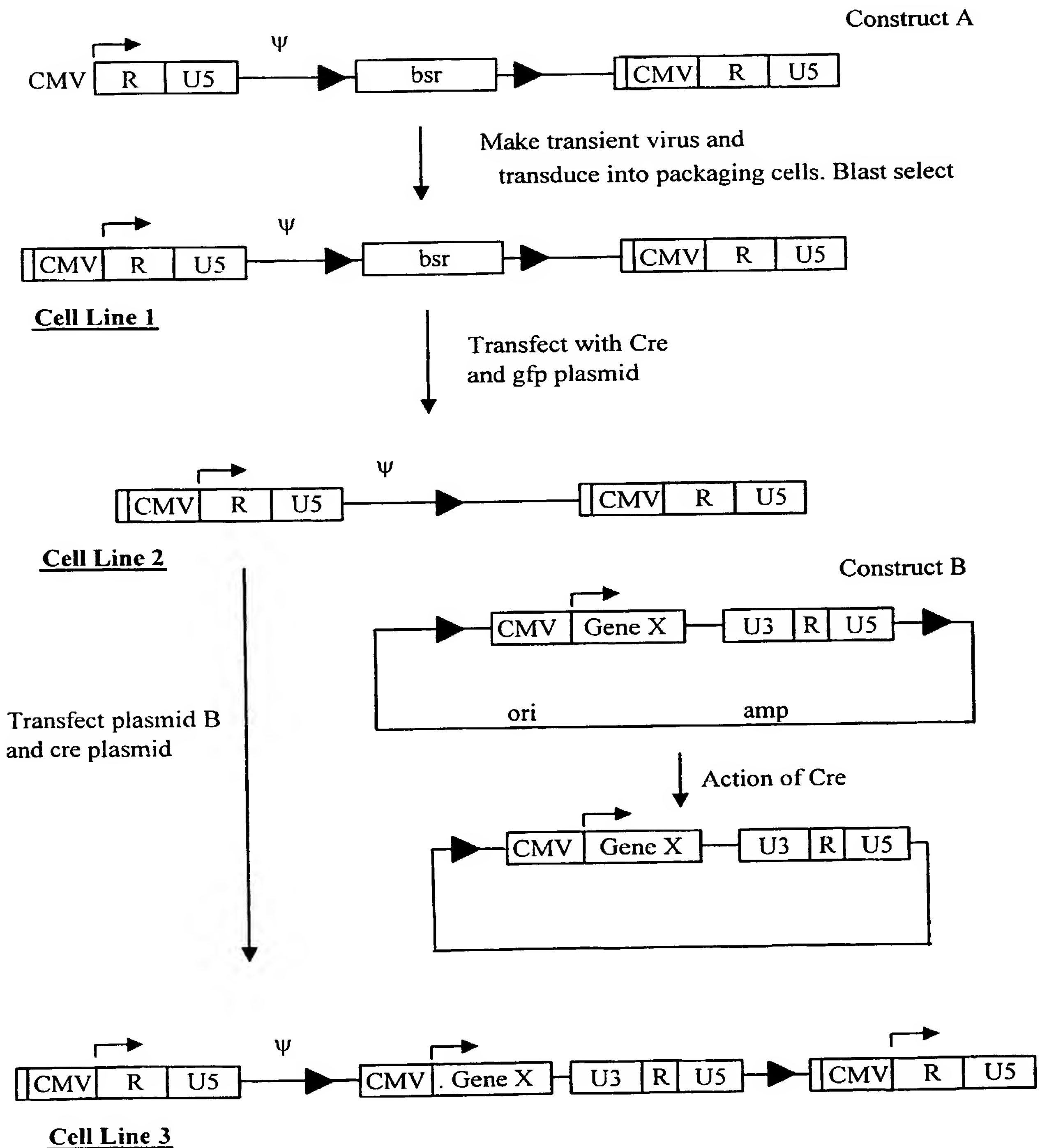


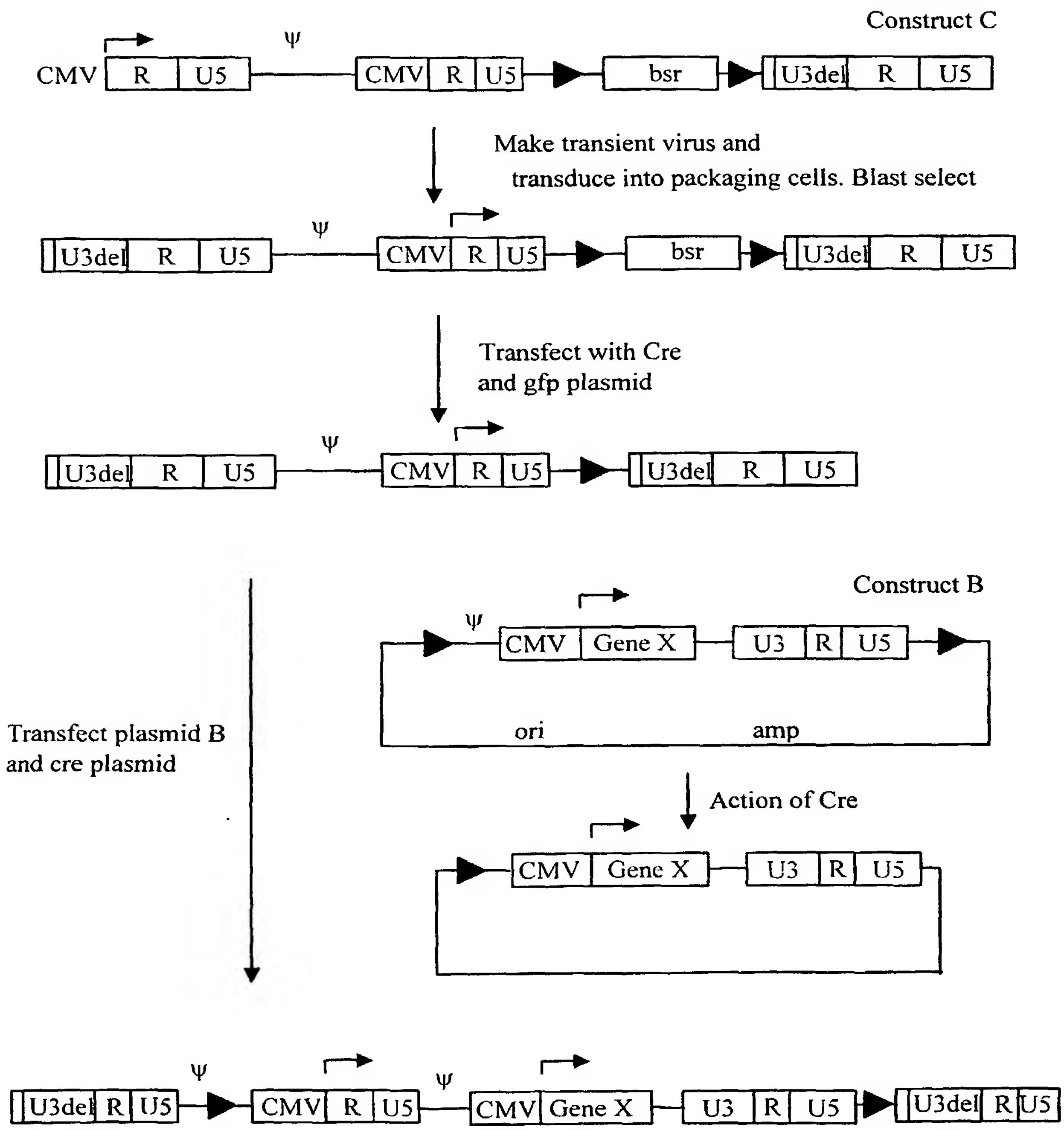
FIG. 2

EIAV-based transduction Cre/loxP system



# FIG. 3

MLV SIN vector approach, with EIAV components in blue



# FIG. 4

MLV-based transduction with HRE 3' LTR using Cre/loxP system

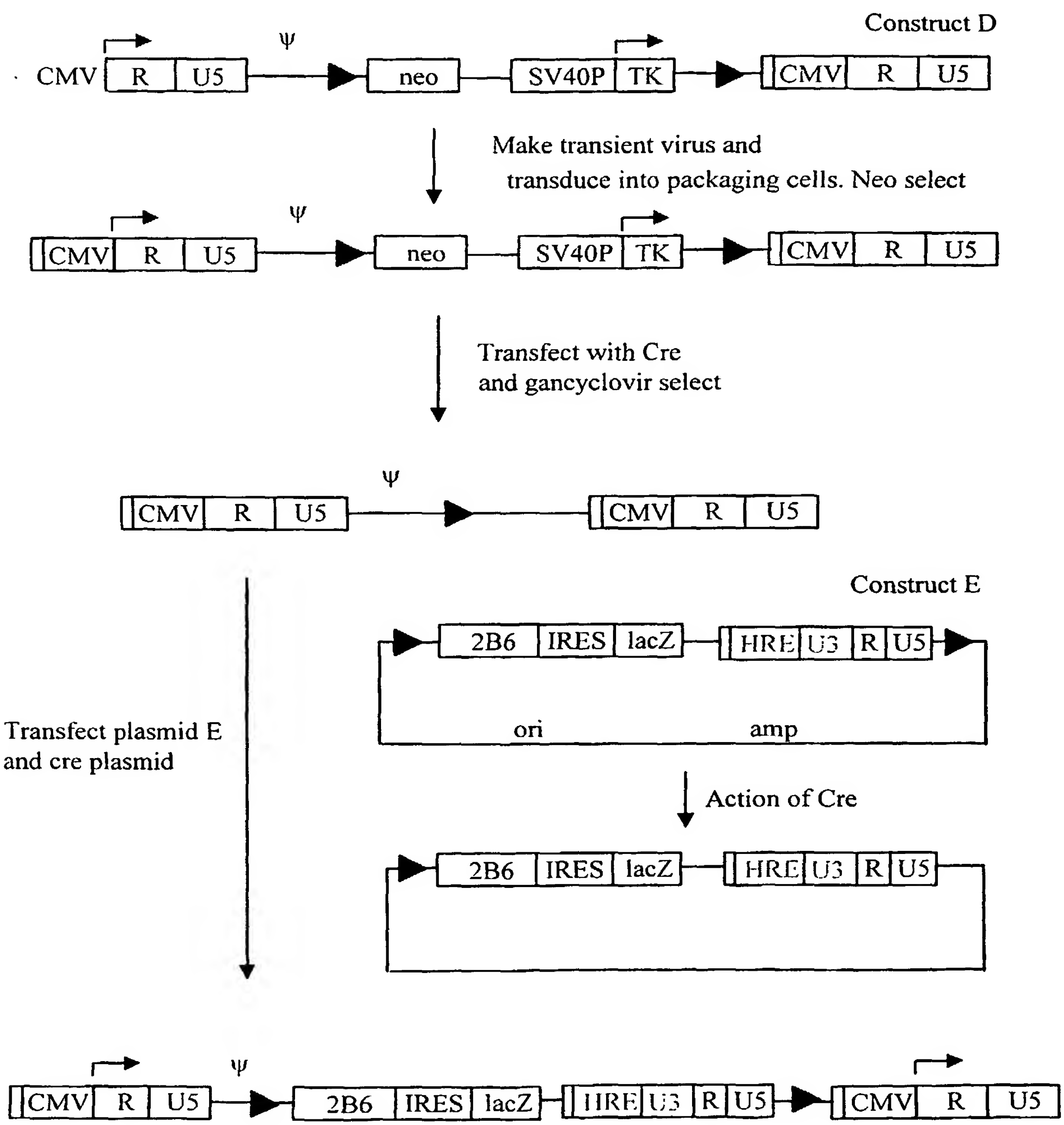
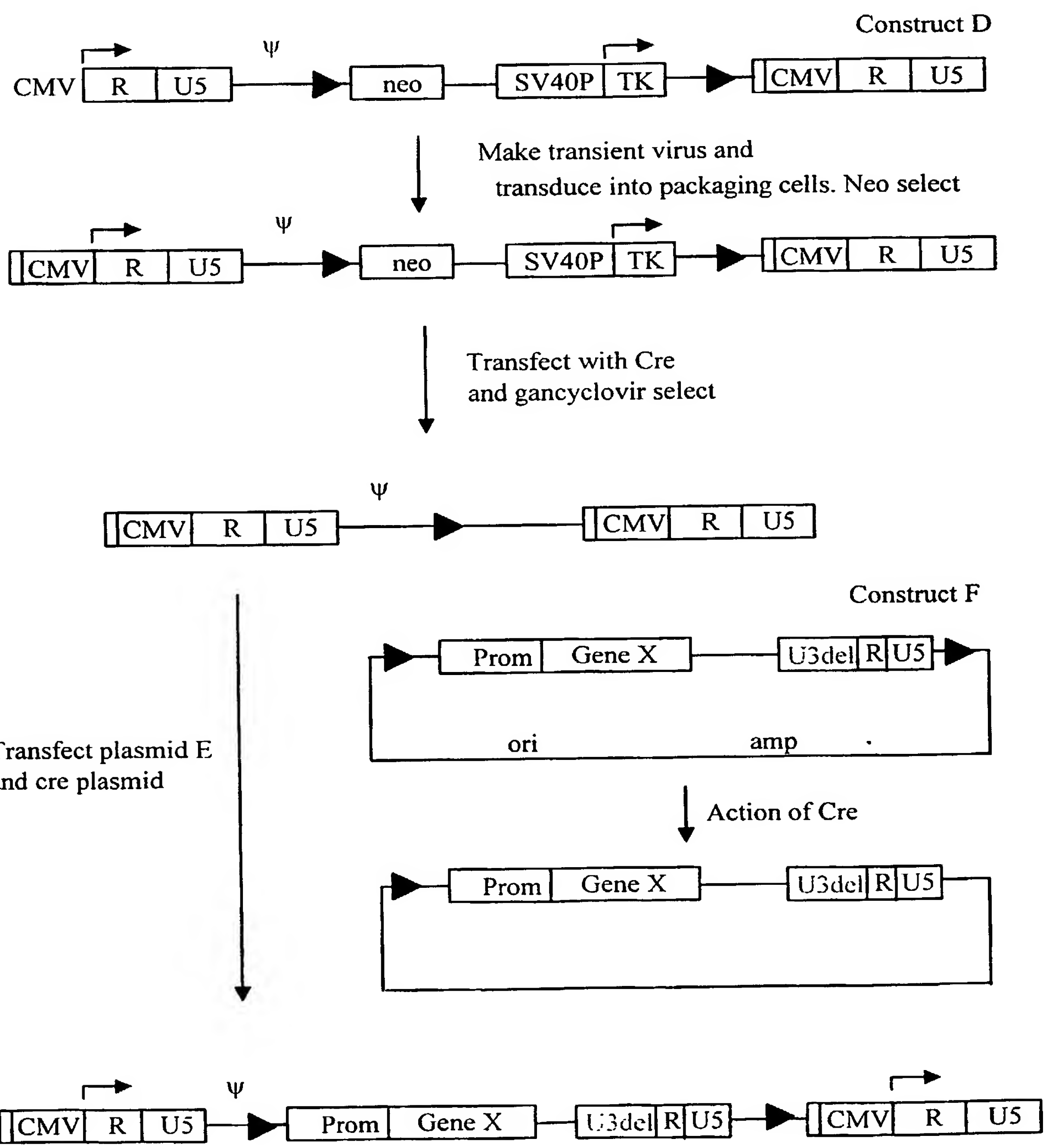


FIG. 5

MLV-based transduction for SIN vector production using Cre/loxP system



# FIG. 6

MLV SIN-vector based transduction system. This general approach can be used with EIAV, HIV or MLV genomes

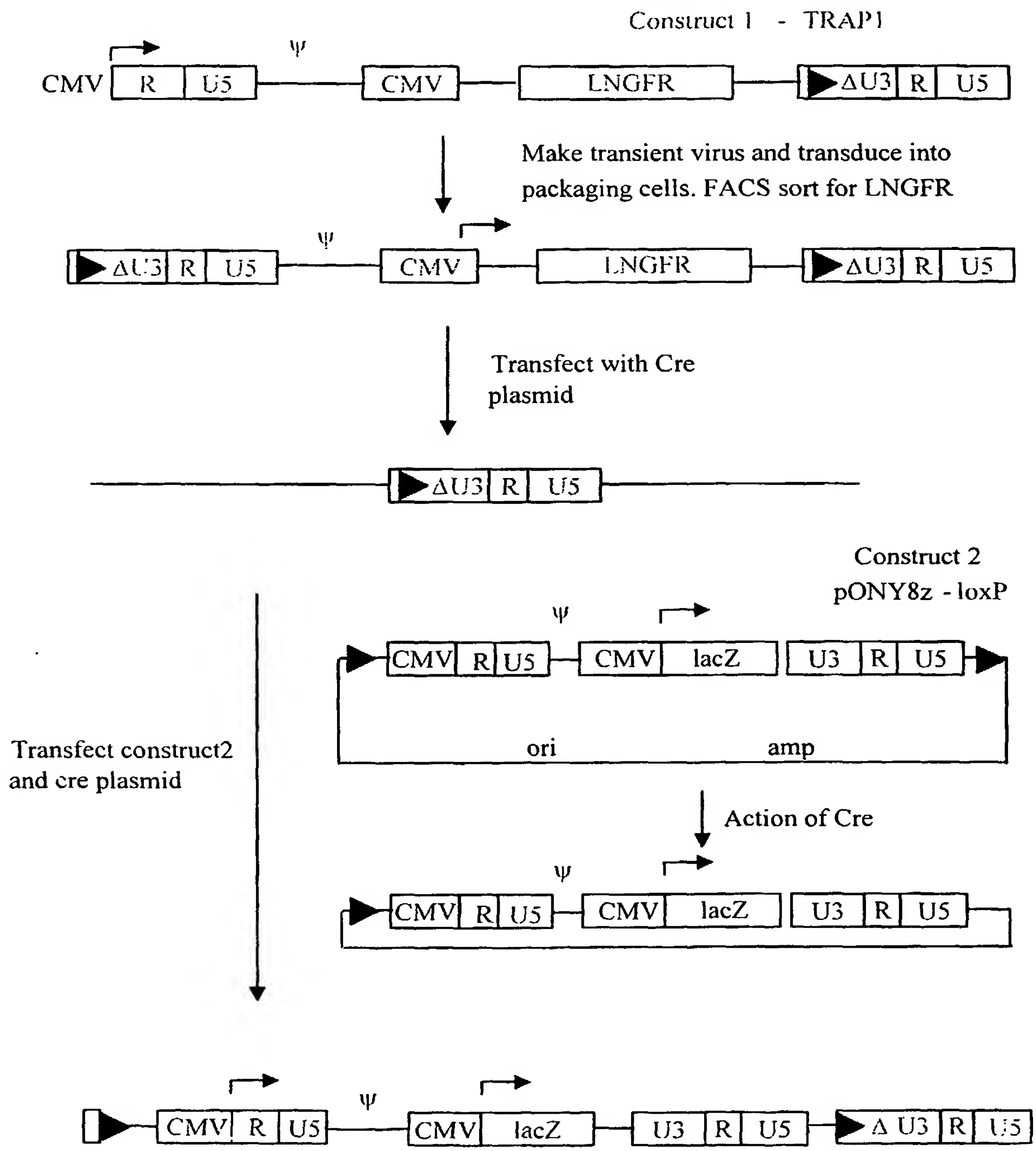


FIG. 7

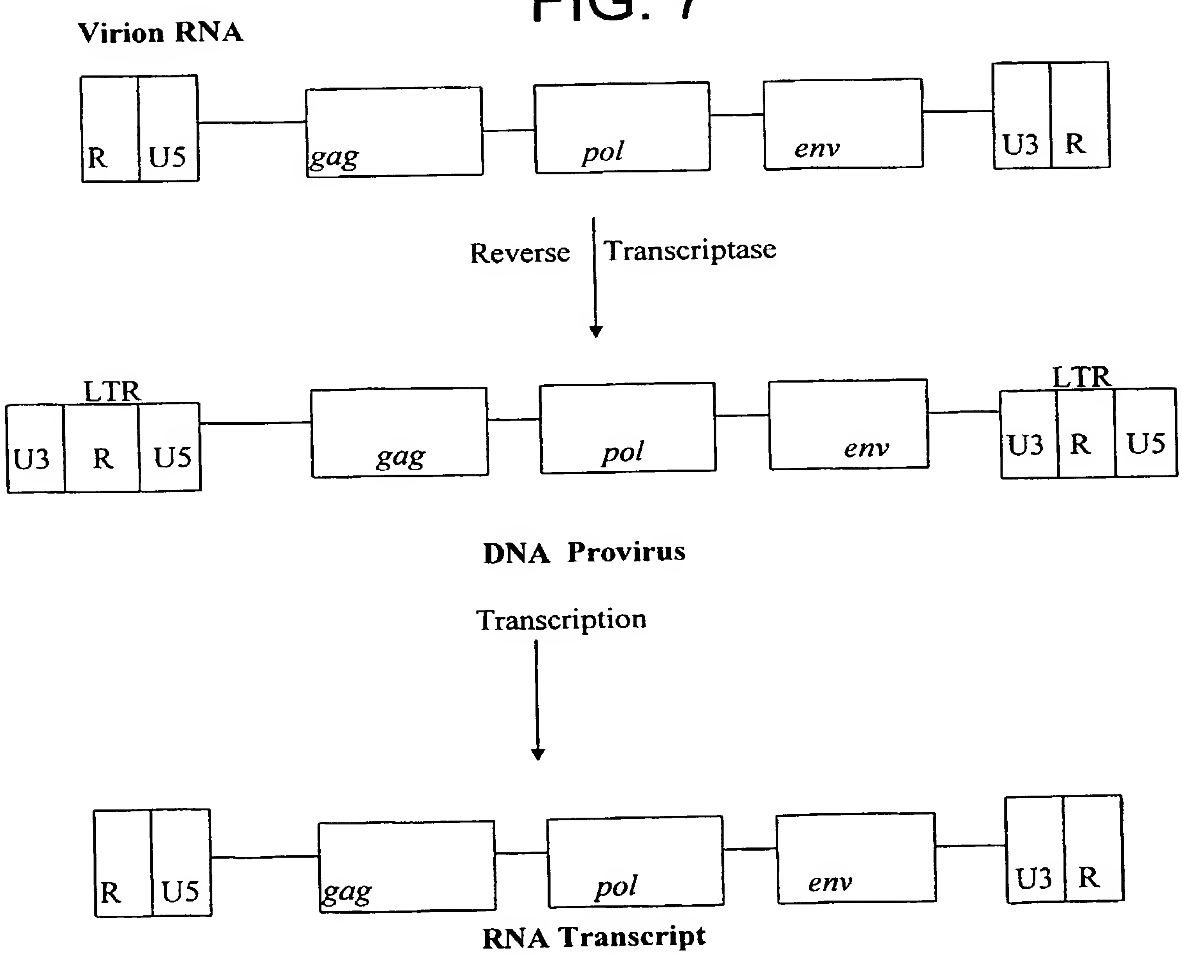

Virion RNA
R
U5
gag
pol
env
U3
R
Reverse
Transcriptase
LTR
U3
R
U5
gag
pol
env
LTR
U3
R
U5
DNA Provirus
Transcription
R
U5
gag
pol
env
U3
R
RNA Transcript

FIG. 8

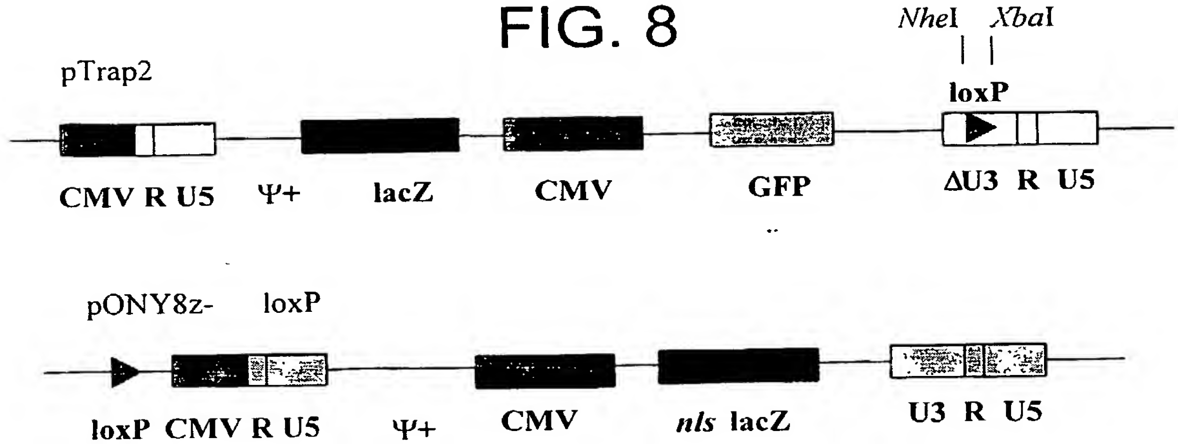
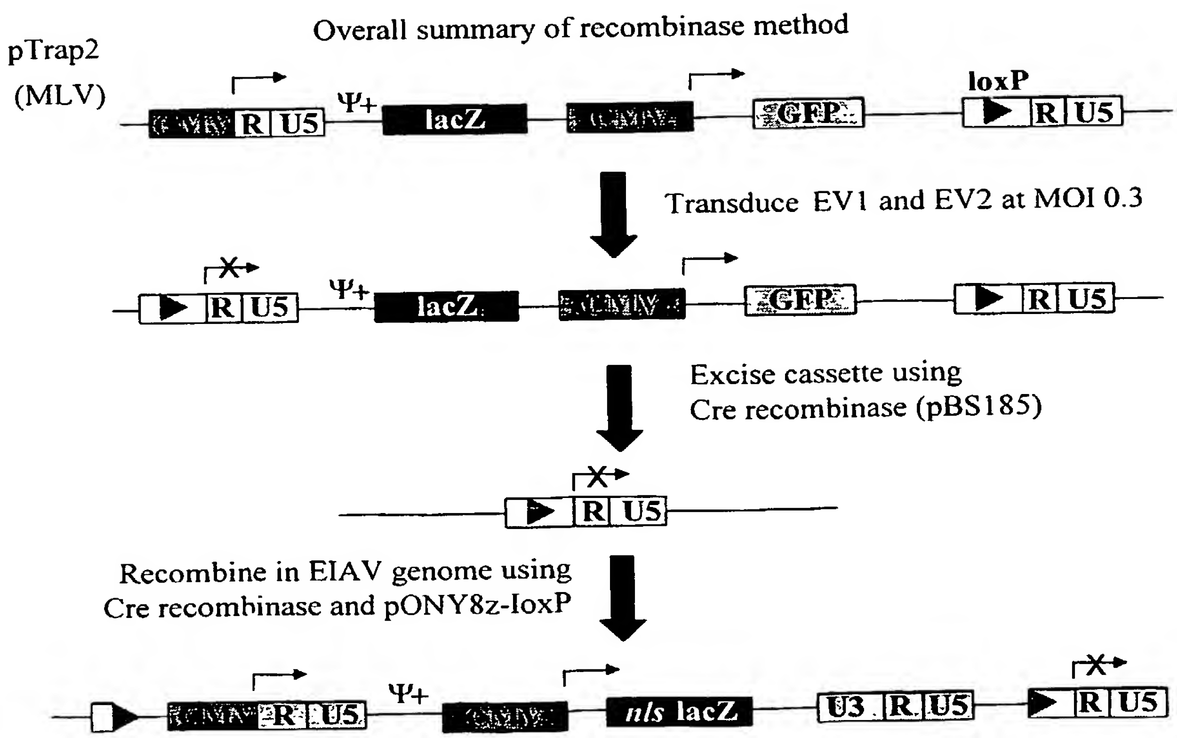

NheI XbaI
pTrap2
loxP
CMV R U5
Ψ+
lacZ
CMV
GFP
ΔU3 R U5
pONY8z- loxP
loxP CMV R U5
Ψ+
CMV
nls lacZ
U3 R U5

FIG. 9


Overall summary of recombinase method
pTrap2
(MLV)
CMV R U5
Ψ+
lacZ
CMV
GFP
loxP
R U5
Transduce EV1 and EV2 at MOI 0.3
Excise cassette using
Cre recombinase (pBS185)
Recombine in EIAV genome using
Cre recombinase and pONY8z-loxP
nls lacZ
U3 R U5

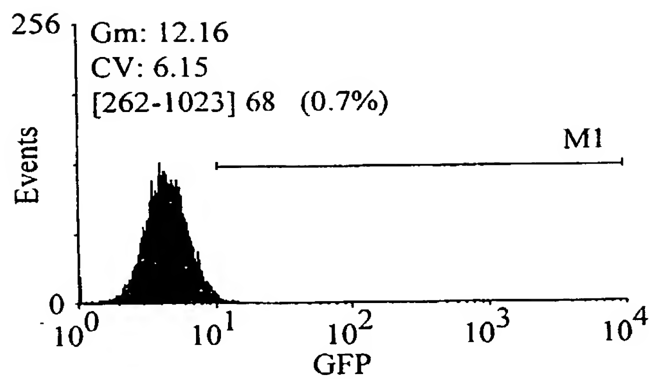



FIG. 10a

FACS analysis of EV1 packaging cells prior to transduction with Trap2 vector

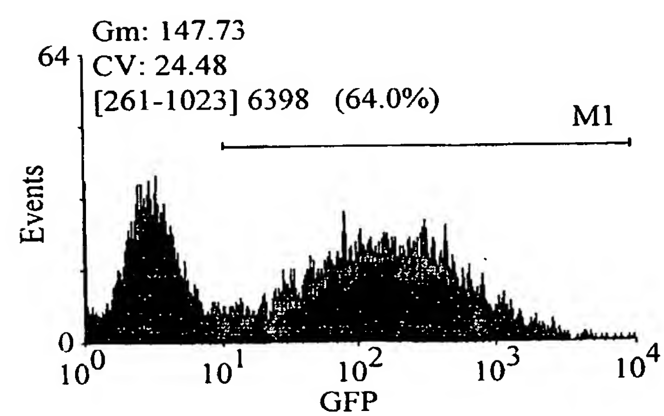
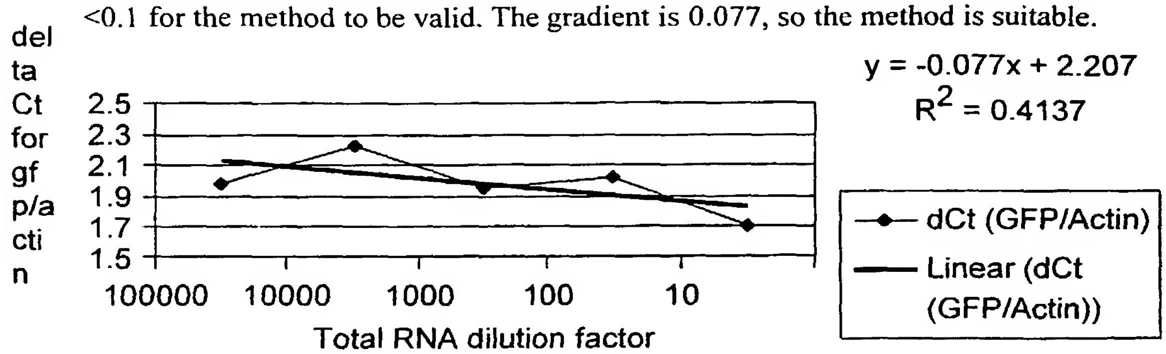



FIG. 10b

FACS analysis of EV1 packaging cell line transduced with Trap2 at an MOI of 0.3. A 5% top-slice of the highest expressers was carried out

# FIG. 11

Validation of the ΔΔCt method for quantitation of GFP mRNA, relative to β-actin. A titration of total RNA from EV1 clone A was used. The difference in Ct values between the two assays is shown on the y-axis. The magnitude of the gradient must be <0.1 for the method to be valid. The gradient is 0.077, so the method is suitable.




# FIG. 12

Quantitation of GFP mRNA relative to control β-actin mRNA. EV2 TD cells are transduced with Trap2 at an MOI of 0.3 and are the calibrator sample with the ratio designated 1.0.

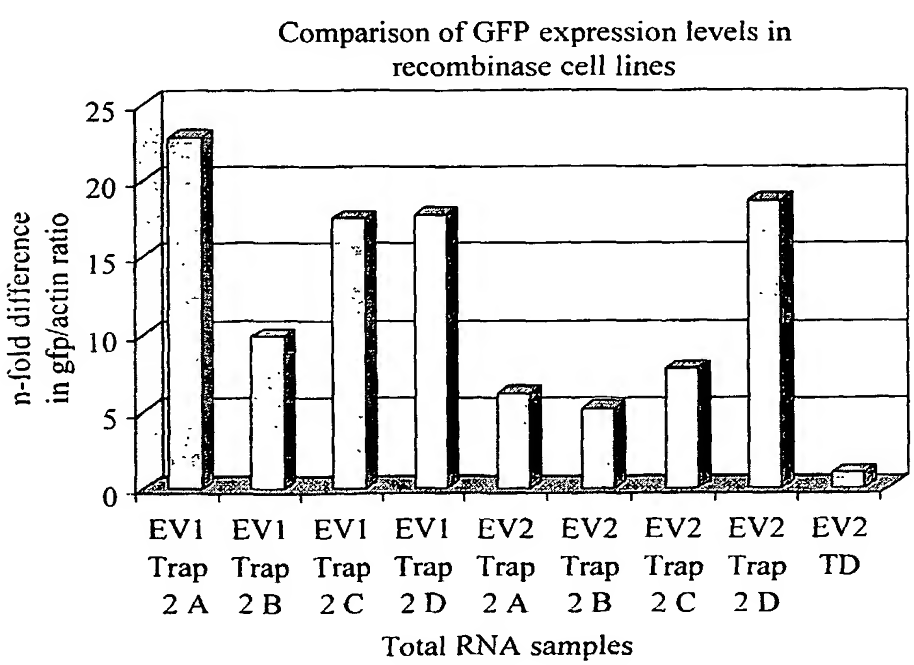
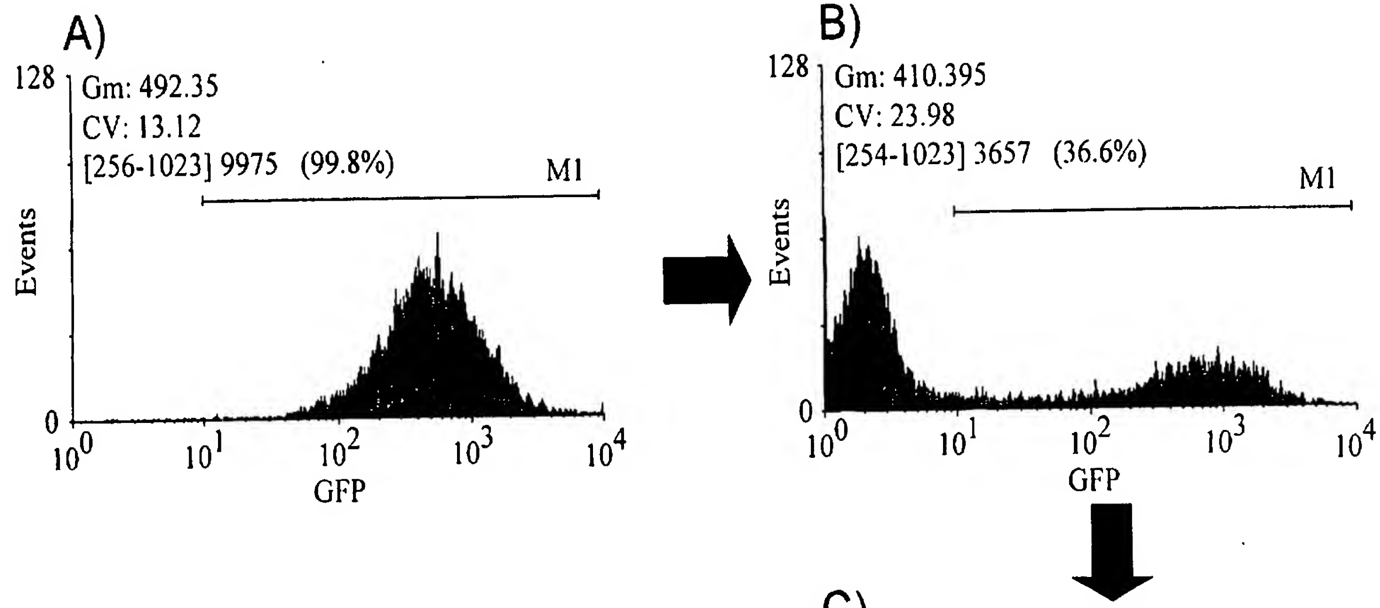
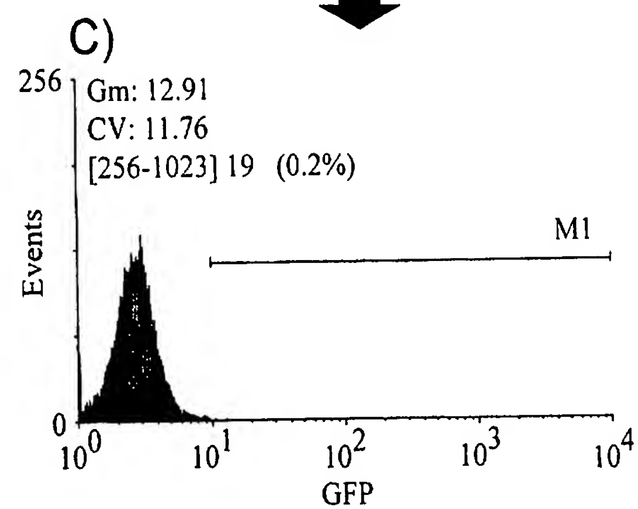


FIG. 13




A) Original GFP expression of the clone.

B) GFP expression 7 days after transfection with Cre recombinase (pBS185). Excision frequency is 64%

C) Recombined clone 4 was identified as being negative for GFP

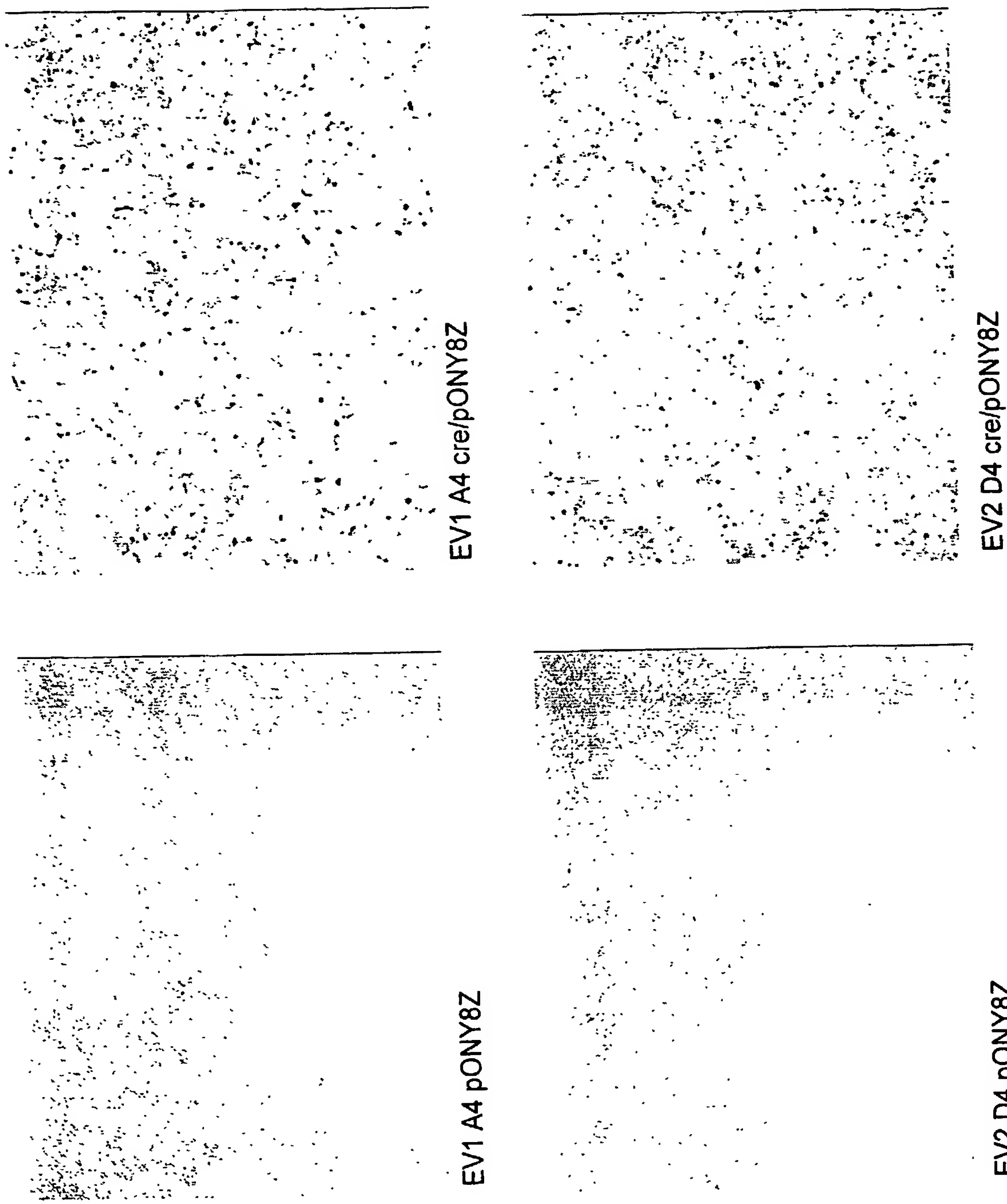

EV1 A4 cre/pONY8Z
EV2 D4 cre/pONY8Z
EV1 A4 pONY8Z
EV2 D4 pONY8Z

FIG. 14

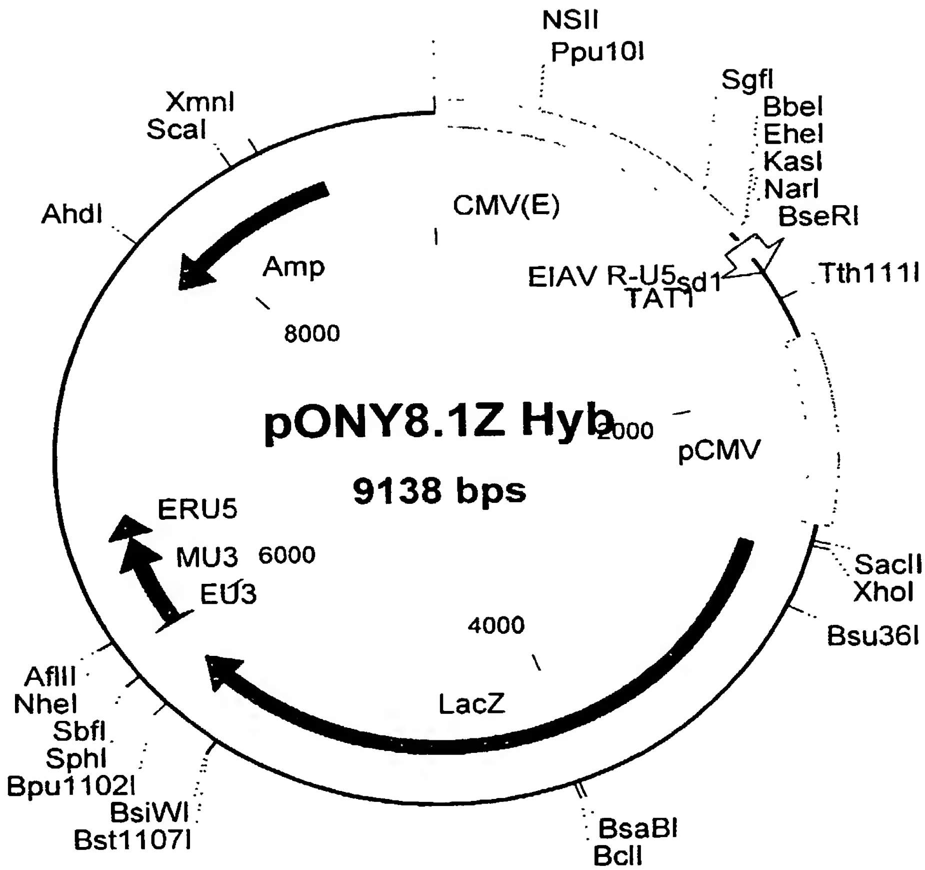

NSII
Ppu10I
SgfI
BbeI
EheI
KasI
NarI
BseRI
XmnI
ScaI
AhdI
CMV(E)
Amp
EIAV R-U5
sd1
TAT1
Tth111I
8000
pONY8.1Z Hyb
9138 bps
2000
pCMV
ERU5
MU3
6000
EU3
SacII
XhoI
Bsu36I
4000
LacZ
AflII
NheI
SbfI
SphI
Bpu1102I
BsiWI
Bst1107I
BsaBI
BclI

FIG. 15

# FIG. 16

Alignment of leader and gag regions present in vectors pONY4Z, 8Z and ATG mutated 8Z vector. The later is referred to as pONY8ZA. The sequence aligned are from the NarI site in the leader to the XbaI site between the EIAV gag sequence and the CMV promoter. Sequences in the leader are shown in italic and a space is present upstream of the position of the gag ATG.

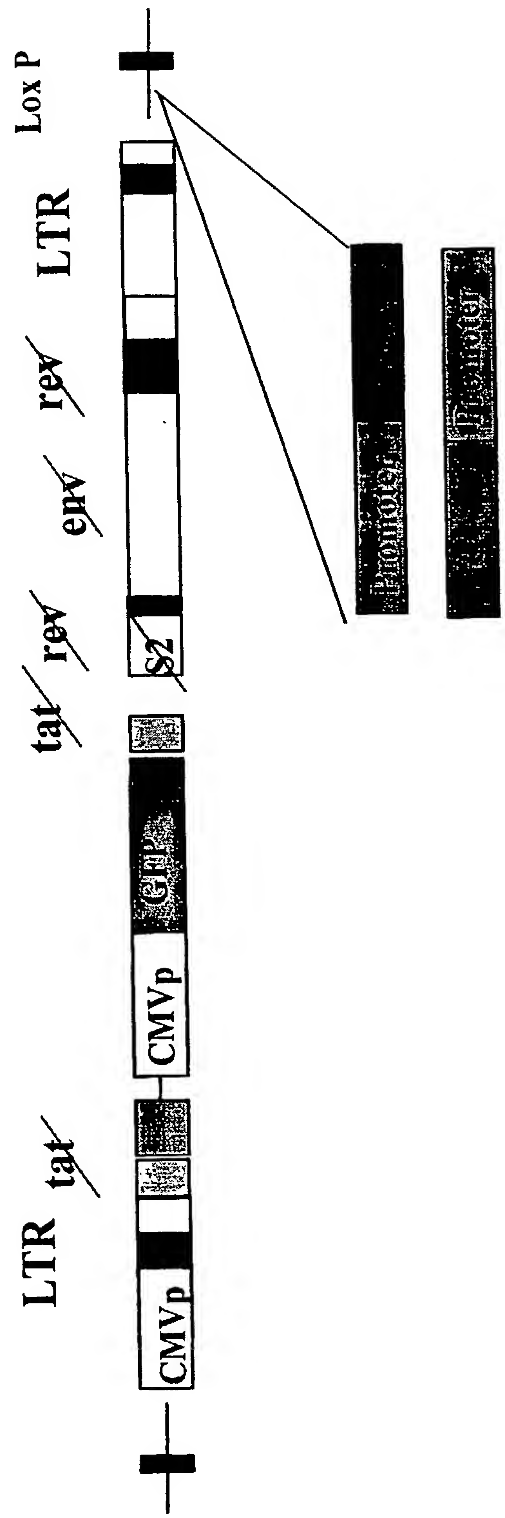
| | |
|---|---|
| 4Z | 1 *cgcccgaacagggacctgagaggggcgcagaccctacctgttgaacctgg* |
| 8Z | 1 *cgcccgaacagggacctgagaggggcgcagaccctacctgttgaacctgg* |
| mutated 8Z | 1 *cgcccgaacagggacctgagaggggcgcagaccctacctgttgaacctgg* |
| | |
| 4Z | 51 *ctgatcgtaggatccccgggacagcagaggagaacttacagaagtcttct* |
| 8Z | 51 *ctgatcgtaggatccccgggacagcagaggagaacttacagaagtcttct* |
| mutated 8Z | 51 *ctgatcgtaggatccccgggacagcagaggagaacttacagaagtcttct* |
| | |
| 4Z | 101 *ggaggtgttcctggccagaacacaggaggacaggtaag*.at-gggagaccc |
| 8Z | 101 *ggaggtgttcctggccagaacacaggaggacaggtaag*.attgggagaccc |
| mutated 8Z | 101 *ggaggtgttcctggccagaacacaggaggacaggtaag*.attgggagaccc |
| | |
| 4Z | 150 tttgacat-ggagcaaggcgctcaagaagttagagaaggtgacggtacaa |
| 8Z | 151 tttgacattggagcaaggcgctcaagaagttagagaaggtgacggtacaa |
| mutated 8Z | 151 tttgacattggagcaaggcgctcaagaagttagagaaggtgacggtacaa |
| | |
| 4Z | 199 gggtctcagaaattaactactggtaactgtaattgggcgctaagtctagt |
| 8Z | 201 gggtctcagaaattaactactggtaactgtaattgggcgctaagtctagt |
| mutated 8Z | 201 gggtctcagaaattaactactggtaactgtaattgggcgctaagtctagt |
| | |
| 4Z | 249 agacttatttcat-gataccaactttgtaaaagaaaaggactggcagctg |
| 8Z | 251 agacttatttcat-gataccaactttgtaaaagaaaaggactggcagctg |
| mutated 8Z | 251 agacttatttcattgataccaactttgtaaaagaaaaggactggcagctg |

| | |
|---|---|
| 4Z | 298 agggat-gtcattccattgctggaagat-gtaactcagacgctgtcagga |
| 8Z | 300 agggat-gtcattccattgctggaagat-gtaactcagacgctgtcagga |
| mutated 8Z | 301 agggattgtcattccattgctggaagattgtaactcagacgctgtcagga |
| | |
| 4Z | 346 caagaaagagaggcctttgaaagaacat-ggtgggcaatttctgctgtaa |
| 8Z | 348 caagaaagagaggcctttgaaagaacat-ggtgggcaatttctgctgtaa |
| mutated 8Z | 351 caagaaagagaggcctttgaaagaacattggtgggcaatttctgctgtaa |
| | |
| 4Z | 395 agat-gggcctccagattaataat-gtagtagat-ggaaaggcatcattc |
| 8Z | 397 agat-gggcctccagattaataat-gtagtagat-ggaaaggcatcattc |
| mutated 8Z | 401 agattgggcctccagattaataattgtagtagattggaaaggcatcattc |
| | |
| 4Z | 442 cagctcctaagagcgaaatat-gaaaagaagactgctaataaaaagcagt |
| 8Z | 444 cagctcctaagagcgaaatat-gaaaagaagactgctaataaaaagcagt |
| mutated 8Z | 451 cagctcctaagagcgaaatattgaaaagaagactgctaataaaaagcagt |
| | |
| 4Z | 491 ctgagccctctgaagaatatct |
| 8Z | 493 ctgagccctctgaagaatatct |
| mutated 8Z | 501 ctgagccctctgaagaatatct |

FIG. 16 CONT'D

FIG. 17

Schematic representation of the structure of pONY 8.3G +/- vector genome plasmids



39

Atty. Dkt. No. 078883-0146

# DECLARATION AND POWER OF ATTORNEY

As a below named inventor, I HEREBY DECLARE:

THAT my residence, post office address, and citizenship are as stated below next to my name;

THAT I believe I am the original, first, and sole inventor (if only one inventor is named below) or an original, first, and joint inventor (if plural inventors are named below or in an attached Declaration) of the subject matter which is claimed and for which a patent is sought on the invention entitled

PRODUCER CELL FOR THE PRODUCTION OF RETROVIRAL VECTORS

(Attorney Docket No. 078883-0146)

the specification of which (check one)

____ is attached hereto.

X was filed on 10/05/2000 as United States Application Number or PCT International Application Number PCT/GB00/03837 (if applicable).

THAT I do not know and do not believe that the same invention was ever known or used by others in the United States of America, or was patented or described in any printed publication in any country, before I (we) invented it;

THAT I do not know and do not believe that the same invention was patented or described in any printed publication in any country, or in public use or on sale in the United States of America, for more than one year prior to the filing date of this United States application;

THAT I do not know and do not believe that the same invention was first patented or made the subject of an inventor's certificate that issued in any country foreign to the United States of America before the filing date of this United States application if the foreign application was filed by me (us), or by my (our) legal representatives or assigns, more than twelve months (six months for design patents) prior to the filing date of this United States application;

THAT I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment specifically referred to above;

THAT I believe that the above-identified specification contains a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the invention, and sets forth the best mode contemplated by me of carrying out the invention; and

THAT I acknowledge the duty to disclose to the U.S. Patent and Trademark Office all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, §1.56.

Atty. Dkt. No. 078883-0146

I HEREBY CLAIM foreign priority benefits under Title 35, United States Code §119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or §365(a) of any PCT international application which designated at least one country other than the United States of America, listed below and have also identified below any foreign application for patent or inventor's certificate or of any PCT international application having a filing date before that of the application on which priority is claimed.

| Prior Foreign Application Number | Country | Foreign Filing Date | Priority Claimed? | Certified Copy Attached? |
|---|---|---|---|---|
| 9923558.2 | Great Britain | 10/05/1999 | YES | |
| | | | | |
| | | | | |

I HEREBY CLAIM the benefit under Title 35, United States Code § 119(e) of any United States provisional application(s) listed below.

| U.S. Provisional Application Number | Filing Date |
|---|---|
| | |
| | |
| | |

I HEREBY CLAIM the benefit under Title 35, United States Code, §120 of any United States application(s), or § 365(c) of any PCT international application designating the United States of America, listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT International application in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose information which is material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which became available between the filing date of the prior application and the national or PCT international filing date of this application.

| U.S. Parent Application Number | PCT Parent Application Number | Parent Filing Date | Parent Patent Number |
|---|---|---|---|
| | | | |
| | | | |
| | | | |

I HEREBY APPOINT the registered attorneys and agents at Customer Number 22428

22428

PATENT TRADEMARK OFFICE

to have full power to prosecute this application and any continuations, divisions, reissues, and reexaminations thereof, to receive the patent, and to transact all business in the United States Patent and Trademark Office connected therewith.

Atty. Dkt. No. 078883-0146

I request that all correspondence be directed to:

Bernhard D. Saxe
FOLEY & LARDNER
Customer Number: 22428

22428
PATENT TRADEMARK OFFICE

Telephone: (202) 672-5427
Facsimile: (202) 672-5399

I UNDERSTAND AND AGREE THAT the foregoing attorneys and agents appointed by me to prosecute this application do not personally represent me or my legal interests, but instead represent the interests of the legal owner(s) of the invention described in this application.

I FURTHER DECLARE THAT all statements made herein of my own knowledge are true, and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

1-W

| | |
|---|---|
| Name of first inventor | Jason SLINGSBY |
| Residence | London, Great Britain |
| Citizenship | Great Britain GBX |
| Post Office Address | 91 Lacy Road, Putney, London SW15 1NR, Great Britain |
| Inventor's signature | Jason H Slingsby |
| Date | 23rd April 2002 |

2-W

| | |
|---|---|
| Name of second inventor | Susan Mary KINGSMAN |
| Residence | Oxford, Great Britain GBX |
| Citizenship | Great Britain |
| Post Office Address | c/o Oxford BioMedica (UK) Ltd.<br>Medwar Centre<br>Robert Robinson Avenue<br>The Oxford Science Park<br>Oxford<br>OX4 4GA<br>Great Britain |
| Inventor's signature | [signature] |
| Date | 19 April 2002 |



002.800057.1

Atty. Dkt. No. 078883-0146

| | |
|---|---|
| 3-W Name of third inventor | Jonathan ROHLL |
| Residence | Oxford, Great Britain GBX |
| Citizenship | Great Britain |
| Post Office Address | c/o Oxford BioMedica (UK) Ltd.<br>Medawar Centre<br>Robert Robinson Avenue<br>The Oxford Science Park<br>Oxford<br>OX4 4GA<br>Great Britain |
| Inventor's signature | J. B. Rohll. |
| 4-W Date | 25th April 2002 |
| Name of fourth inventor | Andrew SLADE |
| Residence | Oxford, Great Britain |
| Citizenship | Great Britain GBX |
| Post Office Address | c/o Oxford BioMedica (UK) Ltd.<br>Medawar Centre<br>Robert Robinson Avenue<br>The Oxford Science Park<br>Oxford<br>OX4 4GA<br>Great Britain |
| Inventor's signature | Andrew Slade |
| Date | 19th April 2002 |



002.800057.1

SEQUENCE LISTING

<110> SLINGSBY, JASON
KINGSMAN, SUSAN
ROHL, JONATHAN
SLADE, ANDREW

<120> PRODUCER CELL FOR THE PRODUCTION OF RETROVIRAL VECTORS

<130> 078883-0146

<140> 10/088,076
<141> 2002-03-20

<150> PCT/GB00/03837
<151> 2000-10-05

<150> GB 9923558.2
<151> 1999-10-05

<160> 65

<170> PatentIn Ver. 2.1

<210> 1
<211> 23
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 1
caaagcatgc ctgcaggaat tcg 23

<210> 2
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 2
gccaaaccta caggtggggt ctttcattat aaaacccctc ataaaaaccc cacag 55

<210> 3
<211> 176
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 3
caaagcatgc ctgcaggaat tcgatatcaa gcttatcgat accgtcgaat tggaagagct 60
ttaaatcctg gcacatctca tgtatcaatg cctcagtatg tttagaaaaa caaggggga 120
actgtggggt ttttatgagg ggttttataa tgaaagaccc cacctgtagg tttggc 176

<210> 4
<211> 55
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 4
ctgtggggtt tttatgaggg gttttataat gaaagacccc acctgtaggt ttggc 55

<210> 5
<211> 56
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artifical Sequence: Primer

<400> 5
gaagggactc agaccgcaga atctgagtgc cccccgagtg aggggttgtg ggctct 56

<210> 6
<211> 508
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 6
ctgtggggtt tttatgaggg gttttataat gaaagacccc acctgtaggt ttggcaagct 60
agcttaagta acgccatttt gcaaggcatg gaaaaataca taactgagaa tagagaagtt 120
cagatcaagg tcaggaacag atggaacagc tgaatatggg ccaaacagga tatctgtggt 180
aagcagttcc tgccccggct cagggccaag aacagatgga acagctgaat atgggccaaa 240
caggatatct gtggtaagca gttcctgccc cggctcaggg ccaagaacag atggtcccca 300
gatgcggtcc agccctcagc agtttctaga gaaccatcag atgtttccag ggtgccccaa 360
ggacctgaaa tgaccctgtg ccttatttga actaaccaat cagttcgctt ctcgcttctg 420
ttcgcgcgct tctgctcccc gagctcaata aaagagccca caacccctca ctcggggggc 480
actcagattc tgcggtctga gtcccttc 508

<210> 7
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 7
gagcgcagcg agtcagtgag cgag 24

<210> 8
<211> 56
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 8
agagcccaca acccctcact cggggggcac tcagattctg cggtctgagt cccttc 56

<210> 9
<211> 419
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 9
agagcccaca acccctcact cggggggcac tcagattctg cggtctgagt cccttctctg 60
ctgggctgaa aaggcctttg taataaatat aattctctac tcagtccctg tctctagttt 120
gtctgttcga gatcctacag agctcatgcc ttggcgtaat catggtcata gctgtttcct 180
gtgtgaaatt gttatccgct cacaattcca cacaacatac gagccggaag cataaagtgt 240
aaagcctggg gtgcctaatg agtgagctaa ctcacattaa ttgcgttgcg ctcactgccc 300
gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg 360
agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc gctgcgctc 419

<210> 10
<211> 9138
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 10
agatcttgaa taataaaatg tgtgtttgtc cgaaatacgc gttttgagat ttctgtcgcc 60
gactaaattc atgtcgcgcg atagtggtgt ttatcgccga tagagatggc gatattggaa 120
aaattgatat ttgaaaatat ggcatattga aaatgtcgcc gatgtgagtt tctgtgtaac 180
tgatatcgcc atttttccaa aagtgatttt tgggcatacg cgatatctgg cgatagcgct 240
tatatcgttt acgggggatg gcgatagacg actttggtga cttgggcgat tctgtgtgtc 300
gcaaatatcg cagtttcgat ataggtgaca gacgatatga ggctatatcg ccgatagagg 360
cgacatcaag ctggcacatg gccaatgcat atcgatctat acattgaatc aatattggcc 420
attagccata ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca 480
tacgttgtat ccatatcgta atatgtacat ttatattggc tcatgtccaa cattaccgcc 540
atgttgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 600
tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 660
gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 720

```
agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt     780
acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg gtaaatggcc     840
cgcctggcat tatgcccagt acatgacctt acgggacttt cctacttggc agtacatcta     900
cgtattagtc atcgctatta ccatggtgat gcggttttgg cagtacacca atgggcgtgg     960
atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt    1020
gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactgcg atcgcccgcc    1080
ccgttgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctcgt    1140
ttagtgaacc gggcactcag attctgcggt ctgagtccct tctctgctgg gctgaaaagg    1200
cctttgtaat aaatataatt ctctactcag tccctgtctc tagtttgtct gttcgagatc    1260
ctacagttgg cgcccgaaca gggacctgag aggggcgcag accctacctg ttgaacctgg    1320
ctgatcgtag gatccccggg acagcagagg agaacttaca gaagtcttct ggaggtgttc    1380
ctggccagaa cacaggagga caggtaagat tgggagaccc tttgacattg gagcaaggcg    1440
ctcaagaagt tagagaaggt gacggtacaa gggtctcaga aattaactac tggtaactgt    1500
aattgggcgc taagtctagt agacttattt catgatacca actttgtaaa agaaaaggac    1560
tggcagctga gggatgtcat tccattgctg gaagatgtaa ctcagacgct gtcaggacaa    1620
gaaagagagg cctttgaaag aacatggtgg gcaatttctg ctgtaaagat gggcctccag    1680
attaataatg tagtagatgg aaaggcatca ttccagctcc taagagcgaa atatgaaaag    1740
aagactgcta ataaaaagca gtctgagccc tctgaagaat atctctagaa ctagtggatc    1800
ccccgggctg caggagtggg gaggcacgat ggccgctttg gtcgaggcgg atccggccat    1860
tagccatatt attcattggt tatatagcat aaatcaatat tggctattgg ccattgcata    1920
cgttgtatcc atatcataat atgtacattt atattggctc atgtccaaca ttaccgccat    1980
gttgacattg attattgact agttattaat agtaatcaat tacggggtca ttagttcata    2040
gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct ggctgaccgc    2100
ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag    2160
ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac    2220
atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg    2280
cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg    2340
tattagtcat cgctattacc atggtgatgc ggttttggca gtacatcaat gggcgtggat    2400
agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt    2460
tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc ccattgacgc    2520
aaatgggcgg taggcatgta cggtgggagg tctatataag cagagctcgt ttagtgaacc    2580
gtcagatcgc ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc    2640
gatccagcct ccgcggcccc aagcttcagc tgctcgagga tctgcggatc cggggaattc    2700
cccagtctca ggatccacca tgggggatcc cgtcgtttta caacgtcgtg actgggaaaa    2760
ccctggcgtt acccaactta atcgccttgc agcacatccc cctttcgcca gctggcgtaa    2820
tagcgaagag gcccgcaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg    2880
gcgctttgcc tggtttccgg caccagaagc ggtgccggaa agctggctgg agtgcgatct    2940
tcctgaggcc gatactgtcg tcgtcccctc aaactggcag atgcacggtt acgatgcgcc    3000
catctacacc aacgtaacct atcccattac ggtcaatccg ccgtttgttc ccacggagaa    3060
tccgacgggt tgttactcgc tcacatttaa tgttgatgaa agctggctac aggaaggcca    3120
gacgcgaatt atttttgatg gcgttaactc ggcgtttcat ctgtggtgca acgggcgctg    3180
ggtcggttac ggccaggaca gtcgtttgcc gtctgaattt gacctgagcg catttttacg    3240
cgccggagaa aaccgcctcg cggtgatggt gctgcgttgg agtgacggca gttatctgga    3300
agatcaggat atgtggcgga tgagcggcat tttccgtgac gtctcgttgc tgcataaacc    3360
gactacacaa atcagcgatt tccatgttgc cactcgcttt aatgatgatt tcagccgcgc    3420
tgtactggag gctgaagttc agatgtgcgg cgagttgcgt gactacctac gggtaacagt    3480
ttctttatgg cagggtgaaa cgcaggtcgc cagcggcacc gcgcctttcg gcggtgaaat    3540
tatcgatgag cgtggtggtt atgccgatcg cgtcacacta cgtctgaacg tcgaaaaccc    3600
gaaactgtgg agcgccgaaa tcccgaatct ctatcgtgcg gtggttgaac tgcacaccgc    3660
cgacggcacg ctgattgaag cagaagcctg cgatgtcggt ttccgcgagg tgcggattga    3720
aaatggtctg ctgctgctga acggcaagcc gttgctgatt cgaggcgtta accgtcacga    3780
gcatcatcct ctgcatggtc aggtcatgga tgagcagacg atggtgcagg atatcctgct    3840
gatgaagcag aacaacttta acgccgtgcg ctgttcgcat tatccgaacc atccgctgtg    3900
gtacacgctg tgcgaccgct acggcctgta tgtggtggat gaagccaata ttgaaaccca    3960
cggcatggtg ccaatgaatc gtctgaccga tgatccgcgc tggctaccgg cgatgagcga    4020
acgcgtaacg cgaatggtgc agcgcgatcg taatcacccg agtgtgatca tctggtcgct    4080
ggggaatgaa tcaggccacg gcgctaatca cgacgcgctg tatcgctgga tcaaatctgt    4140
cgatccttcc cgcccggtgc agtatgaagg cggcggagcc gacaccacgg ccaccgatat    4200
```

tatttgcccg atgtacgcgc gcgtggatga agaccagccc ttcccggctg tgccgaaatg 4260
gtccatcaaa aaatggcttt cgctacctgg agagacgcgc ccgctgatcc tttgcgaata 4320
cgcccacgcg atgggtaaca gtcttggcgg tttcgctaaa tactggcagg cgtttcgtca 4380
gtatccccgt ttacagggcg gcttcgtctg ggactgggtg gatcagtcgc tgattaaata 4440
tgatgaaaac ggcaacccgt ggtcggctta cggcggtgat tttggcgata cgccgaacga 4500
tcgccagttc tgtatgaacg gtctggtctt tgccgaccgc acgccgcatc cagcgctgac 4560
ggaagcaaaa caccagcagc agtttttcca gttccgttta tccgggcaaa ccatcgaagt 4620
gaccagcgaa tacctgttcc gtcatagcga taacgagctc ctgcactgga tggtggcgct 4680
ggatggtaag ccgctggcaa gcggtgaagt gcctctggat gtcgctccac aaggtaaaca 4740
gttgattgaa ctgcctgaac taccgcagcc ggagagcgcc gggcaactct ggctcacagt 4800
acgcgtagtg caaccgaacg cgaccgcatg gtcagaagcc gggcacatca gcgcctggca 4860
gcagtggcgt ctggcggaaa acctcagtgt gacgctcccc gccgcgtccc acgccatccc 4920
gcatctgacc accagcgaaa tggatttttg catcgagctg ggtaataagc gttggcaatt 4980
taaccgccag tcaggctttc tttcacagat gtggattggc gataaaaaac aactgctgac 5040
gccgctgcgc gatcagttca cccgtgcacc gctggataac gacattggcg taagtgaagc 5100
gacccgcatt gaccctaacg cctgggtcga acgctggaag gcggcgggcc attaccaggc 5160
cgaagcagcg ttgttgcagt gcacggcaga tacacttgct gatgcggtgc tgattacgac 5220
cgctcacgcg tggcagcatc aggggaaaac cttatttatc agccggaaaa cctaccggat 5280
tgatggtagt ggtcaaatgg cgattaccgt tgatgttgaa gtggcgagcg atacaccgca 5340
tccggcgcgg attggcctga actgccagct ggcgcaggta gcagagcggg taaactggct 5400
cggattaggg ccgcaagaaa actatcccga ccgccttact gccgcctgtt ttgaccgctg 5460
ggatctgcca ttgtcagaca tgtataccccc gtacgtcttc ccgagcgaaa acggtctgcg 5520
ctgcgggacg cgcgaattga attatggccc acaccagtgg cgcggcgact tccagttcaa 5580
catcagccgc tacagtcaac agcaactgat ggaaaccagc catcgccatc tgctgcacgc 5640
ggaagaaggc acatggctga atatcgacgg tttccatatg ggggattggtg gcgacgactc 5700
ctggagcccg tcagtatcgg cggaattcca gctgagcgcc ggtcgctacc attaccagtt 5760
ggtctggtgt caaaaataat aataaccggg caggggggat ccgcagatcc ggctgtggaa 5820
tgtgtgtcag ttagggtgtg gaaagtcccc aggctcccca gcaggcagaa gtatgcaaag 5880
catgcctgca ggaattcgat atcaagctta tcgataccgt cgaattggaa gagctttaaa 5940
tcctggcaca tctcatgtat caatgcctca gtatgtttag aaaaacaagg ggggaactgt 6000
ggggttttta tgaggggttt tataatgaaa gaccccacct gtaggtttgg caagctagct 6060
taagtaacgc cattttgcaa ggcatggaaa aatacataac tgagaataga gaagttcaga 6120
tcaaggtcag gaacagatgg aacagctgaa tatgggccaa acaggatatc tgtggtaagc 6180
agttcctgcc ccggctcagg gccaagaaca gatggaacag ctgaatatgg gccaaacagg 6240
atatctgtgg taagcagttc ctgccccggc tcagggccaa gaacagatgg tccccagatg 6300
cggtccagcc ctcagcagtt tctagagaac catcagatgt ttccagggtg ccccaaggac 6360
ctgaaatgac cctgtgcctt atttgaacta accaatcagt tcgcttctcg cttctgttcg 6420
cgcgcttctg ctccccgagc tcaataaaag agcccacaac ccctcactcg gggcgcactc 6480
agattctgcg gtctgagtcc cttctctgct gggctgaaaa ggcctttgta ataaatataa 6540
ttctctactc agtccctgtc tctagtttgt ctgttcgaga tcctacagag ctcatgcctt 6600
ggcgtaatca tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca 6660
caacatacga gccggaagca taaagtgtaa agcctggggt gcctaatgag tgagctaact 6720
cacattaatt gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct 6780
gcattaatga atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc 6840
ttcctcgctc actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca 6900
ctcaaaggcg gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg 6960
agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca 7020
taggctccgc ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa 7080
cccgacagga ctataaagat accaggcgtt tccccctgga agctccctcg tgcgctctcc 7140
tgttccgacc ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc 7200
gctttctcat agctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct 7260
gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg 7320
tcttgagtcc aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag 7380
gattagcaga gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta 7440
cggctacact agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg 7500
aaaaagagtt ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt 7560
tgtttgcaag cagcagatta cgcgcagaaa aaaggatctc aagaagatcc tttgatcttt 7620
ttctacgggg tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag 7680

attatcaaaa aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat 7740
ctaaagtata tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc 7800
tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactccccg tcgtgtagat 7860
aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 7920
acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 7980
aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 8040
agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 8100
ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 8160
agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 8220
tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 8280
tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 8340
attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 8400
taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 8460
aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 8520
caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 8580
gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt 8640
cctttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt 8700
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 8760
acctaaattg taagcgttaa tattttgtta aaattcgcgt taaatttttg ttaaatcagc 8820
tcatttttta accaataggc cgaaatcggc aaaatccctt ataaatcaaa agaatagacc 8880
gagatagggt tgagtgttgt tccagtttgg aacaagagtc cactattaaa gaacgtggac 8940
tccaacgtca aagggcgaaa aaccgtctat cagggcgatg gcccactacg tgaaccatca 9000
ccctaatcaa gtttttgggg gtcgaggtgc cgtaaagcac taaatcggaa ccctaaaggg 9060
agcccccgat ttagagcttg acggggaaag ccaacctggc ttatcgaaat taatacgact 9120
cactataggg agaccggc 9138

<210> 11
<211> 56
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 11
ggccatcgtg cctccccact cctgcagtta taaaacccct cataaaaacc ccacag 56

<210> 12
<211> 57
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 12
ctgtggggtt tttatgaggg gttttataaa ctgcaggagt ggggaggcac gatggcc 57

<210> 13
<211> 64
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 13
gaagggactc agaccgcaga atctgagtgc ccggttcact aaacgagctc tgcttatata 60
gacc 64

<210> 14
<211> 64
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 14
ggtctatata agcagagctc gtttagtgaa ccgggcactc agattctgcg gtctgagtcc 60
cttc 64

<210> 15
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 15
cgagatccta cagttggcgc ccgaacag 28

<210> 16
<211> 98
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 16
gagttacaat cttccagcaa tggaatgaca atccctcagc tgccagtcct tttcttttac 60
aaagttggta tcaatgaaat aagtctacta gacttagc 98

<210> 17
<211> 95
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 17
ttccattgct ggaagattgt aactcagacg ctgtcaggac aagaaagaga ggcctttgaa 60
agaacattgg tgggcaattt ctgctgtaaa gattg 95

<210> 18
<211> 98
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 18
caatatttcg ctcttaggag ctggaatgat gcctttccaa tctactacaa ttattaatct 60
ggaggcccaa tctttacagc agaaattgcc caccaatg 98

<210> 19
<211> 83
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 19
ccactagttc tagagatatt cttcagaggg ctcagactgc tttttattag cagtcttctt 60
ttcaatattt cgctcttagg agc 83

<210> 20
<211> 552
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 20
cgagatccta cagttggcgc ccgaacaggg acctgagagg ggcgcagacc ctacctgttg 60
aacctggctg atcgtaggat ccccgggaca gcagaggaga acttacagaa gtcttctgga 120
ggtgttcctg gccagaacac aggaggacag gtaagattgg gagacccttt gacattggag 180
caaggcgctc aagaagttag agaaggtgac ggtacaaggg tctcagaaat taactactgg 240
taactgtaat tgggcgctaa gtctagtaga cttatttcat tgataccaac tttgtaaaag 300
aaaaggactg gcagctgagg gattgtcatt ccattgctgg aagattgtaa ctcagacgct 360
gtcaggacaa gaaagagagg cctttgaaag aacattggtg ggcaatttct gctgtaaaga 420
ttgggcctcc agattaataa ttgtagtaga ttggaaaggc atcattccag ctcctaagag 480
cgaaatattg aaaagaagac tgctaataaa aagcagtctg agccctctga agaatatctc 540
tagaactagt gg 552

<210> 21
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 21
gataacttcg tataatgtat gctatacgaa gttatctgca 40

<210> 22
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 22
gataacttcg tatagcatac attatacgaa gttatctgca 40

<210> 23
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 23
gtgataactt cgtataatgt atgctatacg aagttatcac tac 43

<210> 24
<211> 43
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 24
gtgataactt cgtatagcat acattatacg aagttatcac gta 43

<210> 25
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 25
catgtataac ttcgtataat gtatgctata cgaagttata 40

<210> 26
<211> 40

<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 26
catgtataac ttcgtatagc atacattata cgaagttata 40

<210> 27
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 27
agtaggccgc ctcggccgcc cgggcatca 29

<210> 28
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 28
tgatgcccgg gcggccgagg cggcctact 29

<210> 29
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 29
tagccgagat ctcaaattgc ttagcctgat agcc 34

<210> 30
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 30
tgcgtagcta gcctcccggt ggtgggtcgg tg 32

<210> 31
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 31
agcagtagat ctggggttgg ggttgcgcct tt 32

<210> 32
<211> 31
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 32
cgtcatgcta gcctggggag agaggtcggt g 31

<210> 33
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 33
tacggaagat ctaaatgagt cttcggacct 30

<210> 34
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 34
ctcaacgcta gcgtactcta gccttaagag ctg 33

<210> 35
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 35
taccagagat cttctagagt cgaccaattc tcatg 35

<210> 36
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 36
catcgagcta gcagcttgga ggtgcacacc aatg 34

<210> 37
<211> 32
<212> DNA
<213> Artificial Sequence

<220>
<223> Desription of Artificial Sequence: Primer

<400> 37
gatggtagat ctgcgcagca ccatggcctg aa 32

<210> 38
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 38
ctcgaagcta gcagcttttt gcaaaagcct aggc 34

<210> 39
<211> 515
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 39
gaattcgcca ccatggctga gagcaaggag gccagggatc aagagatgac ctcaaggaag 60
agagcaaaga ggagaagcgc cgcaacgact ggtggaagat cgacccaaag gccccctgga 120
gggggaccag tggtgccgcg tgctgagaca gtccctgccc gaggagaaga ttcctagcca 180
gacctgcatc gccagaagac acctcggccc cggtcccacc cagcacacac cctccagaag 240
ggataggtgg attaggggcc agattttgca agccgaggtc ctccaagaaa ggctggaatg 300

gagaattagg ggcgtgcaac aagccgctaa agagctggga gaggtgaatc gcggcatctg 360
gagggagctc tacttccgcg aggaccagag gggcgatttc tccgcatggg gaggctacca 420
gagggcacaa gaaaggctgt ggggcgagca gagcagcccc cgcgtcttga ggcccggaga 480
ctccaaaaga cgccgcaaac acctgtgaag tcgac 515

<210> 40
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 40
gatcggccgc ctcggcca 18

<210> 41
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 41
gatctggccg aggcggcc 18

<210> 42
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 42
ggccgcctcg gccgta 16

<210> 43
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 43
ggccgaggcg gcctac 16

<210> 44
<211> 38
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<220>
<221> misc_feature
<222> (21)..(22)
<223> The sequence of the EIAV gag/pol ORF is inserted
between these bases

<400> 44
tctagagaat tcgccaccat ggaacccggg gcggccgc 38

<210> 45
<211> 10384
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 45
agatcttgaa taataaaatg tgtgtttgtc cgaaatacgc gttttgagat ttctgtcgcc 60
gactaaattc atgtcgcgcg atagtggtgt ttatcgccga tagagatggc gatattggaa 120
aaattgatat ttgaaaatat ggcatattga aaatgtcgcc gatgtgagtt tctgtgtaac 180
tgatatcgcc atttttccaa aagtgatttt tgggcatacg cgatatctgg cgatagcgct 240
tatatcgttt acgggggatg gcgatagacg actttggtga cttgggcgat tctgtgtgtc 300
gcaaatatcg cagtttcgat ataggtgaca gacgatatga ggctatatcg ccgatagagg 360
cgacatcaag ctggcacatg gccaatgcat atcgatctat acattgaatc aatattggcc 420
attagccata ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca 480
tacgttgtat ccatatcgta atatgtacat ttatattggc tcatgtccaa cattaccgcc 540
atgttgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 600
tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 660
gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 720
agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 780
acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg gtaaatggcc 840
cgcctggcat tatgcccagt acatgacctt acgggacttt cctacttggc agtacatcta 900
cgtattagtc atcgctatta ccatggtgat gcggttttgg cagtacacca atgggcgtgg 960
atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt 1020
gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactgcg atcgcccgcc 1080
ccgttgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctcgt 1140
ttagtgaacc gggcactcag attctgcggt ctgagtccct tctctgctgg gctgaaaagg 1200
cctttgtaat aaatataatt ctctactcag tccctgtctc tagtttgtct gttcgagatc 1260
ctacagttgg cgcccgaaca gggacctgag aggggcgcag accctacctg ttgaacctgg 1320
ctgatcgtag gatccccggg acagcagagg agaacttaca gaagtcttct ggaggtgttc 1380
ctggccagaa cacaggagga caggtaagat tgggagaccc tttgacattg gagcaaggcg 1440
ctcaagaagt tagagaaggt gacggtacaa gggtctcaga aattaactac tggtaactgt 1500
aattgggcgc taagtctagt agacttattt catgatacca actttgtaaa agaaaaggac 1560
tggcagctga gggatgtcat tccattgctg gaagatgtaa ctcagacgct gtcaggacaa 1620
gaaagagagg cctttgaaag aacatggtgg gcaatttctg ctgtaaagat gggcctccag 1680
attaataatg tagtagatgg aaaggcatca ttccagctcc taagagcgaa atatgaaaag 1740

aagactgcta ataaaaagca gtctgagccc tctgaagaat atctctagaa ctagtggatc 1800
ccccgggctg caggagtggg gaggcacgat ggccgctttg gtcgaggcgg atccggccat 1860
tagccatatt attcattggt tatatagcat aaatcaatat tggctattgg ccattgcata 1920
cgttgtatcc atatcataat atgtacattt atattggctc atgtccaaca ttaccgccat 1980
gttgacattg attattgact agttattaat agtaatcaat tacggggtca ttagttcata 2040
gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct ggctgaccgc 2100
ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag 2160
ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac 2220
atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg 2280
cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg 2340
tattagtcat cgctattacc atggtgatgc ggttttggca gtacatcaat gggcgtggat 2400
agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt 2460
tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc ccattgacgc 2520
aaatgggcgg taggcatgta cggtgggagg tctatataag cagagctcgt ttagtgaacc 2580
gtcagatcgc ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc 2640
gatccagcct ccgcggcccc aagcttgttg ggatccaccg gtcgccacca tggtgagcaa 2700
gggcgaggag ctgttcaccg gggtggtgcc catcctggtc gagctggacg gcgacgtaaa 2760
cggccacaag ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg gcaagctgac 2820
cctgaagttc atctgcacca ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac 2880
cctgacctac ggcgtgcagt gcttcagccg ctaccccgac cacatgaagc agcacgactt 2940
cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga 3000
cggcaactac aagacccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat 3060
cgagctgaag ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta 3120
caactacaac agccacaacg tctatatcat ggccgacaag cagaagaacg gcatcaaggt 3180
gaacttcaag atccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca 3240
gcagaacacc cccatcggcg acggccccgt gctgctgccc gacaaccact acctgagcac 3300
ccagtccgcc ctgagcaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt 3360
cgtgaccgcc gccgggatca ctctcggcat ggacgagctg tacaagtaaa gcggccgcga 3420
ctctagagtc gacctgcagg catgcaagct tcagctgctc gaggggggc ccggtaccca 3480
gcttttgttc cctttagtga gggttaattg cgcgggaagt atttatcact aatcaagcac 3540
aagtaataca tgagaaactt ttactacagc aagcacaatc ctccaaaaaa ttttgttttt 3600
acaaaatccc tggtgaacat gattggaagg gacctactag ggtgctgtgg aagggtgatg 3660
gtgcagtagt agttaatgat gaaggaaagg gaataattgc tgtaccatta accaggacta 3720
agttactaat aaaaccaaat tgagtattgt tgcaggaagc aagacccaac taccattgtc 3780
agctgtgttt cctgacctca atatttgtta taaggtttga tatgaatccc aggggaatc 3840
tcaacccta ttacccaaca gtcagaaaaa tctaagtgtg aggagaacac aatgtttcaa 3900
ccttattgtt ataataatga cagtaagaac agcatggcag aatcgaagga agcaagagac 3960
caagaatgaa cctgaaagaa gaatctaaag aagaaaaaag aagaaatgac tggtggaaaa 4020
taggtatgtt tctgttatgc ttagcaggaa ctactggagg aatactttgg tggtatgaag 4080
gactcccaca gcaacattat atagggttgg tggcgatagg gggaagatta aacggatctg 4140
gccaatcaaa tgctatagaa tgctggggtt ccttcccggg gtgtagacca tttcaaaatt 4200
acttcagtta tgagaccaat agaagcatgc atatggataa taatactgct acattattag 4260
aagctttaac caatataact gctctataaa taacaaaaca gaattagaaa catggaagtt 4320
agtaaagact tctggcataa ctcctttacc tatttcttct gaagctaaca ctggactaat 4380
tagacataag agagattttg gtataagtgc aatagtggca gctattgtag ccgctactgc 4440
tattgctgct agcgctacta tgtcttatgt tgctctaact gaggttaaca aaataatgga 4500
agtacaaaat catactttg aggtagaaaa tagtactcta aatggtatgg atttaataga 4560
acgacaaata aagatattat atgctatgat tcttcaaaca catgcagatg ttcaactgtt 4620
aaaggaaaga caacaggtag aggagacatt taatttaatt ggatgtatag aaagaacaca 4680
tgtattttgt catactggtc atccctggaa tatgtcatgg ggacatttaa atgagtcaac 4740
acaatgggat gactgggtaa gcaaaatgga agatttaaat caagagatac taactacact 4800
tcatggagcc aggaacaatt tggcacaatc catgataaca ttcaatacac cagatagtat 4860
agctcaattt ggaaaagacc tttggagtca tattggaaat tggattcctg gattgggagc 4920
ttccattata aaatatatag tgatgttttt gcttatttat ttgttactaa cctcttcgcc 4980
taagatcctc agggccctct ggaaggtgac cagtggtgca gggtcctccg gcagtcgtta 5040
cctgaagaaa aaattccatc acaaacatgc atcgcgagaa gacacctggg accaggccca 5100
acacaacata cacctagcag gcgtgaccgg tggatcaggg gacaaatact acaagcagaa 5160
gtactccagg aacgactgga atggagaatc agaggagtac aacaggcggc caaagagctg 5220

ggtgaagtca atcgaggcat ttggagagag ctatatttcc gagaagacca aaggggagat 5280
ttctcagcct ggggcggcta tcaacgagca caagaacggc tctgggggga acaatcctca 5340
ccaagggtcc ttagacctgg agattcgaag cgaaggagga aacatttatg actgttgcat 5400
taaagcccaa gaaggaactc tcgctatccc ttgctgtgga tttcccttat ggctattttg 5460
gggactagta attatagtag gacgcatagc aggctatgga ttacgtggac tcgctgttat 5520
aataaggatt tgtattagag gcttaaattt gatatttgaa ataatcagaa aaatgcttga 5580
ttatattgga agagctttaa atcctggcac atctcatgta tcaatgcctc agtatgttta 5640
gaaaaacaag gggggaactg tggggttttt atgaggggtt ttataaatga ttataagagt 5700
aaaaagaaag ttgctgatgc tctcataacc ttgtataacc caaaggacta gctcatgttg 5760
ctaggcaact aaaccgcaat aaccgcattt gtgacgcgag ttccccattg gtgacgcgtt 5820
aacttcctgt ttttacagta tataagtgct tgtattctga caattgggca ctcagattct 5880
gcggtctgag tcccttctct gctgggctga aaaggccttt gtaataaata taattctcta 5940
ctcagtccct gtctctagtt tgtctgttcg agatcctaca gagctcatgc cttggcgtaa 6000
tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata 6060
cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta 6120
attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gtgatgcccg 6180
ggcggccgag gcggcctacg tgaaccatca cccaaatcaa gtttttttgcg gtcgaggtgc 6240
cgtaaagctc taaatcggaa ccctaaaggg agcccccgat ttagagcttg acggggaaag 6300
ccggcgaacg tggcgagaaa ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg 6360
gcaagtgtag cggtcacgct gcgcgtaacc accacacccg ccgcgcttaa tgcgccgcta 6420
cagggcgcgt ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg 6480
cctcttcgct attacgccag cccggatcga tccttatcgg attttaccac atttgtagag 6540
gttttacttg ctttaaaaaa cctcccacat ctccccctga acctgaaaca taaaatgaat 6600
gcaattgttg ttgttaactt gtttattgca gcttataatg gttacaaata aagcaatagc 6660
atcacaaatt tcacaaataa agcatttttt tcactgcatt ctagttgtgg tttgtccaaa 6720
ctcatcaatg tatcttatca tgtctgctcg aagcattaac cctcactaaa gggaagcggc 6780
cgcccgggtc gacttcacag gtgtttgcgg cgtcttttgg agtctccggg cctcaagacg 6840
cgggggctgc tctgctcgcc ccacagcctt tcttgtgccc tctggtagcc tccccatgcg 6900
gagaaatcgc ccctctggtc ctcgcggaag tagagctccc tccagatgcc gcgattcacc 6960
tctcccagct ctttagcggc ttgttgcacg cccctaattc tccattccag cctttcttgg 7020
aggacctcgg cttgcaaaat ctggccccta atccacctat cccttctgga gggtgtgtgc 7080
tgggtgggac cggggccgag gtgtcttctg gcgatgcagg tctggctagg aatcttctcc 7140
tcgggcaggg actgtctcag cacgcggcac cactggtccc cctccagggg gccttgtggg 7200
tcgatcttcc accagtcgtt gcggcgcttc tcctctttgc tctcttcctt gaggttcatc 7260
tcttgatccc tggcctcctt gctctcagcc atggtggcga attctcgagg ctagcctccc 7320
ggtggtgggt cggtggtccc tgggcagggg tctccagatc ccggacgagc cccaaatga 7380
aagacccccg agacgggtag tcaatcactc tgaggagacc ctcccaagga acagcgagac 7440
cacgagtcgg atgcaacagc aagaggattt attggataca cgggtacccg ggcgactcag 7500
tctatcggag gactggcgcg ccgagtgagg ggttgtgagc tcttttatag agctcgggaa 7560
gcagaagcgc gcgaacagaa gcgagaagca ggctgattgg ttaattcaaa taaggcacag 7620
ggtcatttca ggtccttggg ggagcctgga aacatctgat gggtcttaag aaactgctga 7680
gggttgggcc atatctgggg accatctgtt cttggccccg ggccggggcc gaaccgcggt 7740
gaccatctgt tcttggcccc gggccggggc cgaaactgct caccgcagat atcctgtttg 7800
gcccaacgtt agctgttttc gtgtacccgc ccttgatctg aacttctcta ttcttggttt 7860
ggtatttttc catgccttgc aaaatggcgt tactgcggct atcaggctaa gcaatttgag 7920
atctggccga ggcggcctac tctgcattaa tgaatcggcc aacgcgcggg gagaggcggt 7980
ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg 8040
ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg 8100
gataacgcag gaaagaacat gtataacttc gtataatgta tgctatacga agttatacat 8160
gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt 8220
ccataggctc cgcccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg 8280
aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 8340
tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 8400
ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa 8460
gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta 8520
tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa 8580
caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa 8640
ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt 8700

```
cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt   8760
ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat   8820
cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat   8880
gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc   8940
aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc   9000
acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta   9060
gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga   9120
cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg   9180
cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc   9240
tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat   9300
cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag   9360
gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat   9420
cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa   9480
ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa   9540
gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga   9600
taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg   9660
gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc   9720
acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg   9780
aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact   9840
cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat   9900
atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt   9960
gccacctaaa ttgtaagcgt taatattttg ttaaaattcg cgttaaattt ttgttaaatc  10020
agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc cttataaatc aaaagaatag  10080
accgagatag ggttgagtgt tgttccagtt tggaacaaga gtccactatt aaagaacgtg  10140
gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg atggcccact acgtgataac  10200
ttcgtataat gtatgctata cgaagttatc actacgtgaa ccatcaccct aatcaagttt  10260
tttggggtcg aggtgccgta aagcactaaa tcggaaccct aaagggagcc cccgatttag  10320
agcttgacgg ggaaagccaa cctggcttat cgaaattaat acgactcact atagggagac  10380
cggc                                                                10384
```

<210> 46
<211> 10384
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 46

```
agatcttgaa taataaaatg tgtgtttgtc cgaaatacgc gttttgagat ttctgtcgcc     60
gactaaattc atgtcgcgcg atagtggtgt ttatcgccga tagagatggc gatattggaa    120
aaattgatat ttgaaaatat ggcatattga aaatgtcgcc gatgtgagtt tctgtgtaac    180
tgatatcgcc atttttccaa aagtgatttt tgggcatacg cgatatctgg cgatagcgct    240
tatatcgttt acgggggatg gcgatagacg actttggtga cttgggcgat tctgtgtgtc    300
gcaaatatcg cagtttcgat ataggtgaca gacgatatga ggctatatcg ccgatagagg    360
cgacatcaag ctggcacatg gccaatgcat atcgatctat acattgaatc aatattggcc    420
attagccata ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca    480
tacgttgtat ccatatcgta atatgtacat ttatattggc tcatgtccaa cattaccgcc    540
atgttgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca    600
tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc    660
gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat    720
agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt    780
acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg gtaaatggcc    840
cgcctggcat tatgcccagt acatgacctt acgggacttt cctacttggc agtacatcta    900
cgtattagtc atcgctatta ccatggtgat gcggttttgg cagtacacca atgggcgtgg    960
atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt   1020
```

gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactgcg atcgcccgcc 1080
ccgttgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctcgt 1140
ttagtgaacc gggcactcag attctgcggt ctgagtccct tctctgctgg gctgaaaagg 1200
cctttgtaat aaatataatt ctctactcag tccctgtctc tagtttgtct gttcgagatc 1260
ctacagttgg cgcccgaaca gggacctgag aggggcgcag accctacctg ttgaacctgg 1320
ctgatcgtag gatccccggg acagcagagg agaacttaca gaagtcttct ggaggtgttc 1380
ctggccagaa cacaggagga caggtaagat tgggagaccc tttgacattg gagcaaggcg 1440
ctcaagaagt tagagaaggt gacggtacaa gggtctcaga aattaactac tggtaactgt 1500
aattgggcgc taagtctagt agacttattt catgatacca actttgtaaa agaaaaggac 1560
tggcagctga gggatgtcat tccattgctg gaagatgtaa ctcagacgct gtcaggacaa 1620
gaaagagagg cctttgaaag aacatggtgg gcaatttctg ctgtaaagat gggcctccag 1680
attaataatg tagtagatgg aaaggcatca ttccagctcc taagagcgaa atatgaaaag 1740
aagactgcta ataaaaagca gtctgagccc tctgaagaat atctctagaa ctagtggatc 1800
ccccgggctg caggagtggg gaggcacgat ggccgctttg gtcgaggcgg atccggccat 1860
tagccatatt attcattggt tatatagcat aaatcaatat tggctattgg ccattgcata 1920
cgttgtatcc atatcataat atgtacattt atattggctc atgtccaaca ttaccgccat 1980
gttgacattg attattgact agttattaat agtaatcaat tacggggtca ttagttcata 2040
gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct ggctgaccgc 2100
ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag 2160
ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac 2220
atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg 2280
cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg 2340
tattagtcat cgctattacc atggtgatgc ggttttggca gtacatcaat gggcgtggat 2400
agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt 2460
tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc ccattgacgc 2520
aaatgggcgg taggcatgta cggtgggagg tctatataag cagagctcgt ttagtgaacc 2580
gtcagatcgc ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc 2640
gatccagcct ccgcggcccc aagcttgttg ggatccaccg gtcgccacca tggtgagcaa 2700
gggcgaggag ctgttcaccg gggtggtgcc catcctggtc gagctggacg gcgacgtaaa 2760
cggccacaag ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg gcaagctgac 2820
cctgaagttc atctgcacca ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac 2880
cctgacctac ggcgtgcagt gcttcagccg ctaccccgac cacatgaagc agcacgactt 2940
cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga 3000
cggcaactac aagacccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat 3060
cgagctgaag ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta 3120
caactacaac agccacaacg tctatatcat ggccgacaag cagaagaacg gcatcaaggt 3180
gaacttcaag atccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca 3240
gcagaacacc cccatcggcg acggccccgt gctgctgccc gacaaccact acctgagcac 3300
ccagtccgcc ctgagcaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt 3360
cgtgaccgcc gccgggatca ctctcggcat ggacgagctg tacaagtaaa gcggccgcga 3420
ctctagagtc gacctgcagg catgcaagct tcagctgctc gagggggggc ccggtaccca 3480
gcttttgttc cctttagtga gggttaattg cgcgggaagt atttatcact aatcaagcac 3540
aagtaataca tgagaaactt ttactacagc aagcacaatc ctccaaaaaa ttttgttttt 3600
acaaaatccc tggtgaacat gattggaagg gacctactag ggtgctgtgg aagggtgatg 3660
gtgcagtagt agttaatgat gaaggaaagg gaataattgc tgtaccatta accaggacta 3720
agttactaat aaaaccaaat tgagtattgt tgcaggaagc aagacccaac taccattgtc 3780
agctgtgttt cctgacctca atatttgtta taaggtttga tatgaatccc agggggaatc 3840
tcaaccccta ttacccaaca gtcagaaaaa tctaagtgtg aggagaacac aatgtttcaa 3900
ccttattgtt ataataatga cagtaagaac agcatggcag aatcgaagga agcaagagac 3960
caagaatgaa cctgaaagaa gaatctaaag aagaaaaaag aagaaatgac tggtggaaaa 4020
taggtatgtt tctgttatgc ttagcaggaa ctactggagg aatactttgg tggtatgaag 4080
gactcccaca gcaacattat atagggttgg tggcgatagg gggaagatta aacggatctg 4140
gccaatcaaa tgctatagaa tgctggggtt ccttcccggg gtgtagacca tttcaaaatt 4200
acttcagtta tgagaccaat agaagcatgc atatggataa taatactgct acattattag 4260
aagctttaac caatataact gctctataaa taacaaaaca gaattagaaa catggaagtt 4320
agtaaagact tctggcataa ctcctttacc tatttcttct gaagctaaca ctggactaat 4380
tagacataag agagattttg gtataagtgc aatagtggca gctattgtag ccgctactgc 4440
tattgctgct agcgctacta tgtcttatgt tgctctaact gaggttaaca aaataatgga 4500

agtacaaaat catacttttg aggtagaaaa tagtactcta aatggtatgg atttaataga 4560
acgacaaata aagatattat atgctatgat tcttcaaaca catgcagatg ttcaactgtt 4620
aaaggaaaga caacaggtag aggagacatt taatttaatt ggatgtatag aaagaacaca 4680
tgtattttgt catactggtc atccctggaa tatgtcatgg ggacatttaa atgagtcaac 4740
acaatgggat gactgggtaa gcaaaatgga agatttaaat caagagatac taactacact 4800
tcatggagcc aggaacaatt tggcacaatc catgataaca ttcaatacac cagatagtat 4860
agctcaattt ggaaaagacc tttggagtca tattggaaat tggattcctg gattgggagc 4920
ttccattata aaatatatag tgatgttttt gcttatttat ttgttactaa cctcttcgcc 4980
taagatcctc agggccctct ggaaggtgac cagtggtgca gggtcctccg gcagtcgtta 5040
cctgaagaaa aaattccatc acaaacatgc atcgcgagaa gacacctggg accaggccca 5100
acacaacata cacctagcag gcgtgaccgg tggatcaggg gacaaatact acaagcagaa 5160
gtactccagg aacgactgga atggagaatc agaggagtac aacaggcggc caaagagctg 5220
ggtgaagtca atcgaggcat ttggagagag ctatatttcc gagaagacca aaggggagat 5280
ttctcagcct ggggcggcta tcaacgagca caagaacggc tctgggggga acaatcctca 5340
ccaagggtcc ttagacctgg agattcgaag cgaaggagga aacatttatg actgttgcat 5400
taaagcccaa gaaggaactc tcgctatccc ttgctgtgga tttcccttat ggctattttg 5460
gggactagta attatagtag gacgcatagc aggctatgga ttacgtggac tcgctgttat 5520
aataaggatt tgtattagag gcttaaattt gatatttgaa ataatcagaa aaatgcttga 5580
ttatattgga agagctttaa atcctggcac atctcatgta tcaatgcctc agtatgttta 5640
gaaaaacaag gggggaactg tggggttttt atgaggggtt ttataaatga ttataagagt 5700
aaaaagaaag ttgctgatgc tctcataacc ttgtataacc caaaggacta gctcatgttg 5760
ctaggcaact aaaccgcaat aaccgcattt gtgacgcgag ttccccattg gtgacgcgtt 5820
aacttcctgt ttttacagta tataagtgct tgtattctga caattgggca ctcagattct 5880
gcggtctgag tcccttctct gctgggctga aaaggccttt gtaataaata taattctcta 5940
ctcagtccct gtctctagtt tgtctgttcg agatcctaca gagctcatgc cttggcgtaa 6000
tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata 6060
cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta 6120
attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gagtaggccg 6180
cctcggccag atctcaaatt gcttagcctg atagccgcag taacgccatt ttgcaaggca 6240
tggaaaaata ccaaaccaag aatagagaag ttcagatcaa gggcgggtac acgaaaacag 6300
ctaacgttgg gccaaacagg atatctgcgg tgagcagttt cggccccggc ccggggccaa 6360
gaacagatgg tcaccgcggt tcggccccgg cccggggcca agaacagatg gtccccagat 6420
atggcccaac cctcagcagt ttcttaagac ccatcagatg tttccaggct cccccaagga 6480
cctgaaatga ccctgtgcct tatttgaatt aaccaatcag cctgcttctc gcttctgttc 6540
gcgcgcttct gcttcccgag ctctataaaa gagctcacaa cccctcactc ggcgcgccag 6600
tcctccgata gactgagtcg cccgggtacc cgtgtatcca ataaatcctc ttgctgttgc 6660
atccgactcg tggtctcgct gttccttggg agggtctcct cagagtgatt gactacccgt 6720
ctcgggggtc tttcatttgg gggctcgtcc gggatctgga gacccctgcc cagggaccac 6780
cgacccacca ccgggaggct agcctcgaga attcgccacc atggctgaga gcaaggaggc 6840
cagggatcaa gagatgaacc tcaaggaaga gagcaaagag gagaagcgcc gcaacgactg 6900
gtggaagatc gacccacaag gcccccctgga gggggaccag tggtgccgcg tgctgagaca 6960
gtccctgccc gaggagaaga ttcctagcca gacctgcatc gccagaagac acctcggccc 7020
cggtcccacc cagcacacac cctccagaag ggataggtgg attaggggcc agattttgca 7080
agccgaggtc ctccaagaaa ggctggaatg gagaattagg ggcgtgcaac aagccgctaa 7140
agagctggga gaggtgaatc gcggcatctg gagggagctc tacttccgcg aggaccagag 7200
gggcgatttc tccgcatggg gaggctacca gagggcacaa gaaaggctgt ggggcgagca 7260
gagcagcccc cgcgtcttga ggcccggaga ctccaaaaga cgccgcaaac acctgtgaag 7320
tcgacccggg cggccgcttc cctttagtga gggttaatgc ttcgagcaga catgataaga 7380
tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaaatg ctttatttgt 7440
gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa acaagttaac 7500
aacaacaatt gcattcattt tatgtttcag gttcaggggg agatgtggga ggttttttaa 7560
agcaagtaaa acctctacaa atgtggtaaa atccgataag gatcgatccg ggctggcgta 7620
atagcgaaga ggcccgcacc gatcgccctt cccaacagtt gcgcagcctg aatggcgaat 7680
ggacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac 7740
cgctacactt gccagcgccc tagcgcccgc tcctttcgct ttcttccctt cctttctcgc 7800
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt 7860
tagagcttta cggcacctcg accgcaaaaa acttgatttg ggtgatggtt cacgtaggcc 7920
gcctcggccg cccgggcatc actgcattaa tgaatcggcc aacgcgcggg gagaggcggt 7980

```
ttgcgtattg ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg   8040
ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg   8100
gataacgcag gaaagaacat gtataacttc gtataatgta tgctatacga agttatacat   8160
gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt   8220
ccataggctc cgcccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg   8280
aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc   8340
tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt   8400
ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa   8460
gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta   8520
tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa   8580
caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa   8640
ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt   8700
cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt   8760
ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat   8820
cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat   8880
gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc   8940
aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc   9000
acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta   9060
gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga   9120
cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg   9180
cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc   9240
tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat   9300
cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag   9360
gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat   9420
cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa   9480
ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa   9540
gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga   9600
taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg   9660
gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc   9720
acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg   9780
aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact   9840
cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat   9900
atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt   9960
gccacctaaa ttgtaagcgt taatattttg ttaaaattcg cgttaaattt ttgttaaatc  10020
agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc cttataaatc aaaagaatag  10080
accgagatag ggttgagtgt tgttccagtt tggaacaaga gtccactatt aaagaacgtg  10140
gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg atggcccact acgtgataac  10200
ttcgtataat gtatgctata cgaagttatc actacgtgaa ccatcaccct aatcaagttt  10260
tttggggtcg aggtgccgta aagcactaaa tcggaaccct aaagggagcc cccgatttag  10320
agcttgacgg ggaaagccaa cctggcttat cgaaattaat acgactcact atagggagac  10380
cggc                                                               10384
```

```
<210> 47
<211> 10292
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Sytnthetic
      oligonucleotide

<400> 47
agatcttgaa taataaaatg tgtgtttgtc cgaaatacgc gttttgagat ttctgtcgcc     60
gactaaattc atgtcgcgcg atagtggtgt ttatcgccga tagagatggc gatattggaa    120
aaattgatat ttgaaaatat ggcatattga aaatgtcgcc gatgtgagtt tctgtgtaac    180
tgatatcgcc atttttccaa aagtgatttt tgggcatacg cgatatctgg cgatagcgct    240
tatatcgttt acgggggatg gcgatagacg actttggtga cttgggcgat tctgtgtgtc    300
```

gcaaatatcg cagtttcgat ataggtgaca gacgatatga ggctatatcg ccgatagagg 360
cgacatcaag ctggcacatg gccaatgcat atcgatctat acattgaatc aatattggcc 420
attagccata ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca 480
tacgttgtat ccatatcgta atatgtacat ttatattggc tcatgtccaa cattaccgcc 540
atgttgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 600
tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 660
gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 720
agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 780
acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg gtaaatggcc 840
cgcctggcat tatgcccagt acatgacctt acgggacttt cctacttggc agtacatcta 900
cgtattagtc atcgctatta ccatggtgat gcggttttgg cagtacacca atgggcgtgg 960
atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt 1020
gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactgcg atcgcccgcc 1080
ccgttgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctcgt 1140
ttagtgaacc gggcactcag attctgcggt ctgagtccct tctctgctgg gctgaaaagg 1200
cctttgtaat aaatataatt ctctactcag tccctgtctc tagtttgtct gttcgagatc 1260
ctacagttgg cgcccgaaca gggacctgag aggggcgcag accctacctg ttgaacctgg 1320
ctgatcgtag gatccccggg acagcagagg agaacttaca gaagtcttct ggaggtgttc 1380
ctggccagaa cacaggagga caggtaagat tgggagaccc tttgacattg gagcaaggcg 1440
ctcaagaagt tagagaaggt gacggtacaa gggtctcaga aattaactac tggtaactgt 1500
aattgggcgc taagtctagt agacttattt catgatacca actttgtaaa agaaaaggac 1560
tggcagctga gggatgtcat tccattgctg gaagatgtaa ctcagacgct gtcaggacaa 1620
gaaagagagg cctttgaaag aacatggtgg gcaatttctg ctgtaaagat gggcctccag 1680
attaataatg tagtagatgg aaaggcatca ttccagctcc taagagcgaa atatgaaaag 1740
aagactgcta ataaaaagca gtctgagccc tctgaagaat atctctagaa ctagtggatc 1800
ccccgggctg caggagtggg gaggcacgat ggccgctttg gtcgaggcgg atccggccat 1860
tagccatatt attcattggt tatatagcat aaatcaatat tggctattgg ccattgcata 1920
cgttgtatcc atatcataat atgtacattt atattggctc atgtccaaca ttaccgccat 1980
gttgacattg attattgact agttattaat agtaatcaat tacggggtca ttagttcata 2040
gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct ggctgaccgc 2100
ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag 2160
ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac 2220
atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg 2280
cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg 2340
tattagtcat cgctattacc atggtgatgc ggttttggca gtacatcaat gggcgtggat 2400
agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt 2460
tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc ccattgacgc 2520
aaatgggcgg taggcatgta cggtgggagg tctatataag cagagctcgt ttagtgaacc 2580
gtcagatcgc ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc 2640
gatccagcct ccgcggcccc aagcttgttg ggatccaccg gtcgccacca tggtgagcaa 2700
gggcgaggag ctgttcaccg gggtggtgcc catcctggtc gagctggacg gcgacgtaaa 2760
cggccacaag ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg gcaagctgac 2820
cctgaagttc atctgcacca ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac 2880
cctgacctac ggcgtgcagt gcttcagccg ctaccccgac cacatgaagc agcacgactt 2940
cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga 3000
cggcaactac aagacccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat 3060
cgagctgaag ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta 3120
caactacaac agccacaacg tctatatcat ggccgacaag cagaagaacg gcatcaaggt 3180
gaacttcaag atccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca 3240
gcagaacacc cccatcggcg acggccccgt gctgctgccc gacaaccact acctgagcac 3300
ccagtccgcc ctgagcaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt 3360
cgtgaccgcc gccgggatca ctctcggcat ggacgagctg tacaagtaaa gcggccgcga 3420
ctctagagtc gacctgcagg catgcaagct tcagctgctc gaggggggc ccggtaccca 3480
gcttttgttc cctttagtga gggttaattg cgcgggaagt atttatcact aatcaagcac 3540
aagtaataca tgagaaactt ttactacagc aagcacaatc ctccaaaaaa ttttgttttt 3600
acaaaatccc tggtgaacat gattggaagg gacctactag ggtgctgtgg aagggtgatg 3660
gtgcagtagt agttaatgat gaaggaaagg gaataattgc tgtaccatta accaggacta 3720
agttactaat aaaaccaaat tgagtattgt tgcaggaagc aagacccaac taccattgtc 3780

```
agctgtgttt cctgacctca atatttgtta taaggtttga tatgaatccc agggggaatc   3840
tcaacccta ttacccaaca gtcagaaaaa tctaagtgtg aggagaacac aatgtttcaa   3900
ccttattgtt ataataatga cagtaagaac agcatggcag aatcgaagga agcaagagac   3960
caagaatgaa cctgaaagaa gaatctaaag aagaaaaaag aagaaatgac tggtggaaaa   4020
taggtatgtt tctgttatgc ttagcaggaa ctactggagg aatactttgg tggtatgaag   4080
gactcccaca gcaacattat atagggttgg tggcgatagg gggaagatta aacggatctg   4140
gccaatcaaa tgctatagaa tgctggggtt ccttcccggg gtgtagacca tttcaaaatt   4200
acttcagtta tgagaccaat agaagcatgc atatggataa taatactgct acattattag   4260
aagctttaac caatataact gctctataaa taacaaaaca gaattagaaa catggaagtt   4320
agtaaagact tctggcataa ctcctttacc tatttcttct gaagctaaca ctggactaat   4380
tagacataag agagattttg gtataagtgc aatagtggca gctattgtag ccgctactgc   4440
tattgctgct agcgctacta tgtcttatgt tgctctaact gaggttaaca aaataatgga   4500
agtacaaaat catacttttg aggtagaaaa tagtactcta aatggtatgg atttaataga   4560
acgacaaata aagatattat atgctatgat tcttcaaaca catgcagatg ttcaactgtt   4620
aaaggaaaga caacaggtag aggagacatt taatttaatt ggatgtatag aaagaacaca   4680
tgtattttgt catactggtc atccctggaa tatgtcatgg ggacatttaa atgagtcaac   4740
acaatgggat gactgggtaa gcaaaatgga agatttaaat caagagatac taactacact   4800
tcatggagcc aggaacaatt tggcacaatc catgataaca ttcaatacac cagatagtat   4860
agctcaattt ggaaaagacc tttggagtca tattggaaat tggattcctg gattgggagc   4920
ttccattata aaatatatag tgatgttttt gcttatttat ttgttactaa cctcttcgcc   4980
taagatcctc agggccctct ggaaggtgac cagtggtgca gggtcctccg gcagtcgtta   5040
cctgaagaaa aaattccatc acaaacatgc atcgcgagaa gacacctggg accaggccca   5100
acacaacata cacctagcag gcgtgaccgg tggatcaggg gacaaatact acaagcagaa   5160
gtactccagg aacgactgga atggagaatc agaggagtac aacaggcggc caaagagctg   5220
ggtgaagtca atcgaggcat ttggagagag ctatatttcc gagaagacca aaggggagat   5280
ttctcagcct ggggcggcta tcaacgagca caagaacggc tctgggggga acaatcctca   5340
ccaagggtcc ttagacctgg agattcgaag cgaaggagga aacatttatg actgttgcat   5400
taaagcccaa gaaggaactc tcgctatccc ttgctgtgga tttcccttat ggctattttg   5460
gggactagta attatagtag gacgcatagc aggctatgga ttacgtggac tcgctgttat   5520
aataaggatt tgtattagag gcttaaattt gatatttgaa ataatcagaa aaatgcttga   5580
ttatattgga agagctttaa atcctggcac atctcatgta tcaatgcctc agtatgttta   5640
gaaaaacaag gggggaactg tggggttttt atgaggggtt ttataaatga ttataagagt   5700
aaaaagaaag ttgctgatgc tctcataacc ttgtataacc caaaggacta gctcatgttg   5760
ctaggcaact aaaccgcaat aaccgcattt gtgacgcgag ttccccattg gtgacgcgtt   5820
aacttcctgt ttttacagta tataagtgct tgtattctga caattgggca ctcagattct   5880
gcggtctgag tcccttctct gctgggctga aaaggccttt gtaataaata taattctcta   5940
ctcagtccct gtctctagtt tgtctgttcg agatcctaca gagctcatgc cttggcgtaa   6000
tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata   6060
cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta   6120
attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gtgatgcccg   6180
ggcggccgag gcggcctacg tgaaccatca cccaaatcaa gttttttgcg gtcgaggtgc   6240
cgtaaagctc taaatcggaa ccctaaaggg agcccccgat ttagagcttg acggggaaag   6300
ccggcgaacg tggcgagaaa ggaagggaag aaagcgaaag gagcgggcgc tagggcgctg   6360
gcaagtgtag cggtcacgct gcgcgtaacc accacacccg ccgcgcttaa tgcgccgcta   6420
cagggcgcgt ccattcgcca ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg   6480
cctcttcgct attacgccag cccggatcga tccttatcgg attttaccac atttgtagag   6540
gttttacttg ctttaaaaaa cctcccacat ctccccctga acctgaaaca taaaatgaat   6600
gcaattgttg ttgttaactt gtttattgca gcttataatg gttacaaata aagcaatagc   6660
atcacaaatt tcacaaataa agcatttttt tcactgcatt ctagttgtgg tttgtccaaa   6720
ctcatcaatg tatcttatca tgtctgctcg aagcattaac cctcactaaa gggaagcggc   6780
cgcccgggtc gacttcacag gtgtttgcgg cgtcttttgg agtctccggg cctcaagacg   6840
cgggggctgc tctgctcgcc ccacagcctt tcttgtgccc tctggtagcc tccccatgcg   6900
gagaaatcgc ccctctggtc ctcgcggaag tagagctccc tccagatgcc gcgattcacc   6960
tctcccagct ctttagcggc ttgttgcacg cccctaattc tccattccag cctttcttgg   7020
aggacctcgg cttgcaaaat ctggccccta atccacctat cccttctgga gggtgtgtgc   7080
tgggtgggac cggggccgag gtgtcttctg gcgatgcagg tctggctagg aatcttctcc   7140
tcgggcaggg actgtctcag cacgcggcac cactggtccc cctccagggg gccttgtggg   7200
tcgatcttcc accagtcgtt gcggcgcttc tcctctttgc tctcttcctt gaggttcatc   7260
```

tcttgatccc tggcctcctt gctctcagcc atggtggcga attctcgagg ctagcctggg 7320
gagagaggtc ggtgattcgg tcaacgaggg agccgactgc cgacgtgcgc tccggaggct 7380
tgcagaatgc ggaacaccgc gcgggcagga acagggccca cactaccgcc ccacaccccg 7440
cctcccgcac cgccccttcc cggccgctgc tctcggcgcg ccccgctgag cagccgctat 7500
tggccacagc ccatcgcggt cggcgcgctg ccattgctcc ctggcgctgt ccgtctgcga 7560
gggtactagt gagacgtgcg gcttccgttt gtcacgtccg gcacgccgcg aaccgcaagg 7620
aaccttcccg acttaggggc ggagcaggaa gcgtcgccgg ggggcccaca agggtagcgg 7680
cgaagatccg ggtgacgctg cgaacggacg tgaagaatgt gcgagaccca gggtcggcgc 7740
cgctgcgttt cccggaacca cgcccagagc agccgcgtcc ctgcgcaaac ccagggctgc 7800
cttggaaaag gcgcaacccc aaccccagat ctggccgagg cggcctactc tgcattaatg 7860
aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct 7920
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 7980
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt ataacttcgt 8040
ataatgtatg ctatacgaag ttatacatgt gagcaaaagg ccagcaaaag gccaggaacc 8100
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca 8160
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 8220
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 8280
tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 8340
tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc 8400
ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact 8460
tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg 8520
ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggta 8580
tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca 8640
aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa 8700
aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg 8760
aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc 8820
ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg 8880
acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat 8940
ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg 9000
gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa 9060
taaaccagcc agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca 9120
tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc 9180
gcaacgttgt tgccattgct acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 9240
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 9300
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 9360
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 9420
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 9480
gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag 9540
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 9600
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 9660
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9720
cgacacggaa atgttgaata ctcatactct tcctttttca atattattga agcatttatc 9780
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9840
gggttccgcg cacatttccc cgaaaagtgc cacctaaatt gtaagcgtta atattttgtt 9900
aaaattcgcg ttaaattttt gttaaatcag ctcatttttt aaccaatagg ccgaaatcgg 9960
caaaatccct tataaatcaa aagaatagac cgagataggg ttgagtgttg ttccagtttg 10020
gaacaagagt ccactattaa agaacgtgga ctccaacgtc aaagggcgaa aaaccgtcta 10080
tcagggcgat ggcccactac gtgataactt cgtataatgt atgctatacg aagttatcac 10140
tacgtgaacc atcaccctaa tcaagttttt tggggtcgag gtgccgtaaa gcactaaatc 10200
ggaaccctaa agggagcccc cgatttagag cttgacgggg aaagccaacc tggcttatcg 10260
aaattaatac gactcactat agggagaccg gc 10292

<210> 48
<211> 10292
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 48
agatcttgaa taataaaatg tgtgtttgtc cgaaatacgc gttttgagat ttctgtcgcc 60
gactaaattc atgtcgcgcg atagtggtgt ttatcgccga tagagatggc gatattggaa 120
aaattgatat ttgaaaatat ggcatattga aaatgtcgcc gatgtgagtt tctgtgtaac 180
tgatatcgcc atttttccaa aagtgatttt tgggcatacg cgatatctgg cgatagcgct 240
tatatcgttt acgggggatg gcgatagacg actttggtga cttgggcgat tctgtgtgtc 300
gcaaatatcg cagtttcgat ataggtgaca gacgatatga ggctatatcg ccgatagagg 360
cgacatcaag ctggcacatg gccaatgcat atcgatctat acattgaatc aatattggcc 420
attagccata ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca 480
tacgttgtat ccatatcgta atatgtacat ttatattggc tcatgtccaa cattaccgcc 540
atgttgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 600
tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 660
gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 720
agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 780
acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg gtaaatggcc 840
cgcctggcat tatgcccagt acatgacctt acgggacttt cctacttggc agtacatcta 900
cgtattagtc atcgctatta ccatggtgat gcggttttgg cagtacacca atgggcgtgg 960
atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt 1020
gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactgcg atcgcccgcc 1080
ccgttgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctcgt 1140
ttagtgaacc gggcactcag attctgcggt ctgagtccct tctctgctgg gctgaaaagg 1200
cctttgtaat aaatataatt ctctactcag tccctgtctc tagtttgtct gttcgagatc 1260
ctacagttgg cgcccgaaca gggacctgag aggggcgcag accctacctg ttgaacctgg 1320
ctgatcgtag gatccccggg acagcagagg agaacttaca gaagtcttct ggaggtgttc 1380
ctggccagaa cacaggagga caggtaagat tgggagaccc tttgacattg gagcaaggcg 1440
ctcaagaagt tagagaaggt gacggtacaa gggtctcaga aattaactac tggtaactgt 1500
aattgggcgc taagtctagt agacttattt catgatacca actttgtaaa agaaaaggac 1560
tggcagctga gggatgtcat tccattgctg gaagatgtaa ctcagacgct gtcaggacaa 1620
gaaagagagg cctttgaaag aacatggtgg gcaatttctg ctgtaaagat gggcctccag 1680
attaataatg tagtagatgg aaaggcatca ttccagctcc taagagcgaa atatgaaaag 1740
aagactgcta ataaaaagca gtctgagccc tctgaagaat atctctagaa ctagtggatc 1800
ccccgggctg caggagtggg gaggcacgat ggccgctttg gtcgaggcgg atccggccat 1860
tagccatatt attcattggt tatatagcat aaatcaatat tggctattgg ccattgcata 1920
cgttgtatcc atatcataat atgtacattt atattggctc atgtccaaca ttaccgccat 1980
gttgacattg attattgact agttattaat agtaatcaat tacggggtca ttagttcata 2040
gcccatatat ggagttccgc gttacataac ttacggtaaa tggcccgcct ggctgaccgc 2100
ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt tcccatagta acgccaatag 2160
ggactttcca ttgacgtcaa tgggtggagt atttacggta aactgcccac ttggcagtac 2220
atcaagtgta tcatatgcca agtacgcccc ctattgacgt caatgacggt aaatggcccg 2280
cctggcatta tgcccagtac atgaccttat gggactttcc tacttggcag tacatctacg 2340
tattagtcat cgctattacc atggtgatgc ggttttggca gtacatcaat gggcgtggat 2400
agcggtttga ctcacgggga tttccaagtc tccaccccat tgacgtcaat gggagtttgt 2460
tttggcacca aaatcaacgg gactttccaa aatgtcgtaa caactccgcc ccattgacgc 2520
aaatgggcgg taggcatgta cggtgggagg tctatataag cagagctcgt ttagtgaacc 2580
gtcagatcgc ctggagacgc catccacgct gttttgacct ccatagaaga caccgggacc 2640
gatccagcct ccgcggcccc aagcttgttg ggatccaccg gtcgccacca tggtgagcaa 2700
gggcgaggag ctgttcaccg gggtggtgcc catcctggtc gagctggacg gcgacgtaaa 2760
cggccacaag ttcagcgtgt ccggcgaggg cgagggcgat gccacctacg gcaagctgac 2820
cctgaagttc atctgcacca ccggcaagct gcccgtgccc tggcccaccc tcgtgaccac 2880
cctgacctac ggcgtgcagt gcttcagccg ctaccccgac cacatgaagc agcacgactt 2940
cttcaagtcc gccatgcccg aaggctacgt ccaggagcgc accatcttct tcaaggacga 3000
cggcaactac aagacccgcg ccgaggtgaa gttcgagggc gacaccctgg tgaaccgcat 3060
cgagctgaag ggcatcgact tcaaggagga cggcaacatc ctggggcaca agctggagta 3120
caactacaac agccacaacg tctatatcat ggccgacaag cagaagaacg gcatcaaggt 3180

gaacttcaag atccgccaca acatcgagga cggcagcgtg cagctcgccg accactacca 3240
gcagaacacc cccatcggcg acggccccgt gctgctgccc gacaaccact acctgagcac 3300
ccagtccgcc ctgagcaaag accccaacga gaagcgcgat cacatggtcc tgctggagtt 3360
cgtgaccgcc gccgggatca ctctcggcat ggacgagctg tacaagtaaa gcggccgcga 3420
ctctagagtc gacctgcagg catgcaagct tcagctgctc gaggggggc ccggtaccca 3480
gcttttgttc cctttagtga gggttaattg cgcgggaagt atttatcact aatcaagcac 3540
aagtaataca tgagaaactt ttactacagc aagcacaatc ctccaaaaaa ttttgttttt 3600
acaaaatccc tggtgaacat gattggaagg gacctactag ggtgctgtgg aagggtgatg 3660
gtgcagtagt agttaatgat gaaggaaagg gaataattgc tgtaccatta accaggacta 3720
agttactaat aaaaccaaat tgagtattgt tgcaggaagc aagacccaac taccattgtc 3780
agctgtgttt cctgacctca atatttgtta taaggtttga tatgaatccc agggggaatc 3840
tcaaccccta ttacccaaca gtcagaaaaa tctaagtgtg aggagaacac aatgtttcaa 3900
ccttattgtt ataataatga cagtaagaac agcatggcag aatcgaagga agcaagagac 3960
caagaatgaa cctgaaagaa gaatctaaag aagaaaaag aagaaatgac tggtggaaaa 4020
taggtatgtt tctgttatgc ttagcaggaa ctactggagg aatactttgg tggtatgaag 4080
gactcccaca gcaacattat atagggttgg tggcgatagg gggaagatta aacggatctg 4140
gccaatcaaa tgctatagaa tgctggggtt ccttcccggg gtgtagacca tttcaaaatt 4200
acttcagtta tgagaccaat agaagcatgc atatggataa taatactgct acattattag 4260
aagctttaac caatataact gctctataaa taacaaaaca gaattagaaa catggaagtt 4320
agtaaagact tctggcataa ctcctttacc tatttcttct gaagctaaca ctggactaat 4380
tagacataag agagattttg gtataagtgc aatagtggca gctattgtag ccgctactgc 4440
tattgctgct agcgctacta tgtcttatgt tgctctaact gaggttaaca aaataatgga 4500
agtacaaaat catactttg aggtagaaaa tagtactcta aatggtatgg atttaataga 4560
acgacaaata aagatattat atgctatgat tcttcaaaca catgcagatg ttcaactgtt 4620
aaaggaaaga caacaggtag aggagacatt taatttaatt ggatgtatag aaagaacaca 4680
tgtattttgt catactggtc atccctggaa tatgtcatgg ggacatttaa atgagtcaac 4740
acaatgggat gactgggtaa gcaaaatgga agatttaaat caagagatac taactacact 4800
tcatggagcc aggaacaatt tggcacaatc catgataaca ttcaatacac cagatagtat 4860
agctcaattt ggaaaagacc tttggagtca tattggaaat tggattcctg gattgggagc 4920
ttccattata aaatatatag tgatgttttt gcttatttat ttgttactaa cctcttcgcc 4980
taagatcctc agggccctct ggaaggtgac cagtggtgca gggtcctccg gcagtcgtta 5040
cctgaagaaa aaattccatc acaaacatgc atcgcgagaa gacacctggg accaggccca 5100
acacaacata cacctagcag gcgtgaccgg tggatcaggg gacaaatact acaagcagaa 5160
gtactccagg aacgactgga atggagaatc agaggagtac aacaggcggc caaagagctg 5220
ggtgaagtca atcgaggcat ttggagagag ctatatttcc gagaagacca aaggggagat 5280
ttctcagcct ggggcggcta tcaacgagca caagaacggc tctgggggga acaatcctca 5340
ccaagggtcc ttagacctgg agattcgaag cgaaggagga aacatttatg actgttgcat 5400
taaagcccaa gaaggaactc tcgctatccc ttgctgtgga tttcccttat ggctattttg 5460
gggactagta attatagtag gacgcatagc aggctatgga ttacgtggac tcgctgttat 5520
aataaggatt tgtattagag gcttaaattt gatatttgaa ataatcagaa aaatgcttga 5580
ttatattgga agagctttaa atcctggcac atctcatgta tcaatgcctc agtatgttta 5640
gaaaaacaag gggggaactg tggggttttt atgaggggtt ttataaatga ttataagagt 5700
aaaaagaaag ttgctgatgc tctcataacc ttgtataacc caaaggacta gctcatgttg 5760
ctaggcaact aaaccgcaat aaccgcattt gtgacgcgag ttccccattg gtgacgcgtt 5820
aacttcctgt ttttacagta tataagtgct tgtattctga caattgggca ctcagattct 5880
gcggtctgag tcccttctct gctgggctga aaaggccttt gtaataaata taattctcta 5940
ctcagtccct gtctctagtt tgtctgttcg agatcctaca gagctcatgc cttggcgtaa 6000
tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata 6060
cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta 6120
attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca gagtaggccg 6180
cctcggccag atctggggtt ggggttgcgc cttttccaag gcagccctgg gtttgcgcag 6240
ggacgcggct gctctgggcg tggttccggg aaacgcagcg gcgccgaccc tgggtctcgc 6300
acattcttca cgtccgttcg cagcgtcacc cggatcttcg ccgctaccct tgtgggcccc 6360
ccggcgacgc ttcctgctcc gcccctaagt cgggaaggtt ccttgcggtt cgcggcgtgc 6420
cggacgtgac aaacggaagc cgcacgtctc actagtaccc tcgcagacgg acagcgccag 6480
ggagcaatgg cagcgcgccg accgcgatgg gctgtggcca atagcggctg ctcagcgggg 6540
cgcgccgaga gcagcggccg ggaaggggcg gtgcgggagg cggggtgtgg ggcggtagtg 6600
tgggccctgt tcctgcccgc gcggtgttcc gcattctgca agcctccgga gcgcacgtcg 6660

gcagtcggct ccctcgttga ccgaatcacc gacctctctc cccaggctag cctcgagaat 6720
tcgccaccat ggctgagagc aaggaggcca gggatcaaga gatgaacctc aaggaagaga 6780
gcaaagagga gaagcgccgc aacgactggt ggaagatcga cccacaaggc cccctggagg 6840
gggaccagtg gtgccgcgtg ctgagacagt ccctgcccga ggagaagatt cctagccaga 6900
cctgcatcgc cagaagacac ctcggccccg gtcccaccca gcacacaccc tccagaaggg 6960
ataggtggat taggggccag atttgcaag ccgaggtcct ccaagaaagg ctggaatgga 7020
gaattagggg cgtgcaacaa gccgctaaag agctgggaga ggtgaatcgc ggcatctgga 7080
gggagctcta cttccgcgag gaccagaggg gcgatttctc cgcatgggga ggctaccaga 7140
gggcacaaga aaggctgtgg ggcgagcaga gcagcccccg cgtcttgagg cccggagact 7200
ccaaaagacg ccgcaaacac ctgtgaagtc gacccgggcg gccgcttccc tttagtgagg 7260
gttaatgctt cgagcagaca tgataagata cattgatgag tttggacaaa ccacaactag 7320
aatgcagtga aaaaaatgct ttatttgtga aatttgtgat gctattgctt tatttgtaac 7380
cattataagc tgcaataaac aagttaacaa caacaattgc attcatttta tgtttcaggt 7440
tcagggggag atgtgggagg ttttttaaag caagtaaaac ctctacaaat gtggtaaaat 7500
ccgataagga tcgatccggg ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc 7560
caacagttgc gcagcctgaa tggcgaatgg acgcgccctg tagcggcgca ttaagcgcgg 7620
cgggtgtggt ggttacgcgc agcgtgaccg ctacacttgc cagcgcccta gcgcccgctc 7680
ctttcgcttt cttcccttcc tttctcgcca cgttcgccgg ctttccccgt caagctctaa 7740
atcgggggct ccctttaggg ttccgattta gagctttacg gcacctcgac cgcaaaaaac 7800
ttgatttggg tgatggttca cgtaggccgc ctcggccgcc cgggcatcac tgcattaatg 7860
aatcggccaa cgcgcgggga gaggcggttt gcgtattggg cgctcttccg cttcctcgct 7920
cactgactcg ctgcgctcgg tcgttcggct gcggcgagcg gtatcagctc actcaaaggc 7980
ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt ataacttcgt 8040
ataatgtatg ctatacgaag ttatacatgt gagcaaaagg ccagcaaaag gccaggaacc 8100
gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg cccccctgac gagcatcaca 8160
aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt 8220
ttccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc 8280
tgtccgcctt tctcccttcg ggaagcgtgg cgctttctca tagctcacgc tgtaggtatc 8340
tcagttcggt gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc 8400
ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact 8460
tatcgccact ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg 8520
ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggta 8580
tctgcgctct gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca 8640
aacaaaccac cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa 8700
aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg 8760
aaaactcacg ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc 8820
ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg 8880
acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat 8940
ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc ttaccatctg 9000
gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa 9060
taaaccagcc agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca 9120
tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc 9180
gcaacgttgt tgccattgct acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt 9240
cattcagctc cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgcaaaa 9300
aagcggttag ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat 9360
cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct 9420
tttctgtgac tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga 9480
gttgctcttg cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag 9540
tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga 9600
gatccagttc gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca 9660
ccagcgtttc tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg 9720
cgacacggaa atgttgaata ctcatactct tcctttttca atattattga agcatttatc 9780
agggttattg tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag 9840
gggttccgcg cacatttccc cgaaaagtgc cacctaaatt gtaagcgtta atattttgtt 9900
aaaattcgcg ttaaattttt gttaaatcag ctcatttttt aaccaatagg ccgaaatcgg 9960
caaaatccct tataaatcaa aagaatagac cgagataggg ttgagtgttg ttccagtttg 10020
gaacaagagt ccactattaa agaacgtgga ctccaacgtc aaagggcgaa aaaccgtcta 10080
tcagggcgat ggcccactac gtgataactt cgtataatgt atgctatacg aagttatcac 10140

tacgtgaacc atcaccctaa tcaagttttt tggggtcgag gtgccgtaaa gcactaaatc 10200
ggaaccctaa agggagcccc cgatttagag cttgacgggg aaagccaacc tggcttatcg 10260
aaattaatac gactcactat agggagaccg gc 10292

<210> 49
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 49
ctagcataac ttcgtataat gtatgctata cgaagttatt 40

<210> 50
<211> 40
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 50
ctagaataac ttcgtatagc atacattata cgaagttatg 40

<210> 51
<211> 12206
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic oligonucleotide

<400> 51
agatctcccg atcccctatg gtcgactctc agtacaatct gctctgatgc cgcatagtta 60
agccagtatc tgctccctgc ttgtgtgttg gaggtcgctg agtagtgcgc gagcaaaatt 120
taagctacaa caaggcaagg cttgaccgac aattgcatga agaatctgct tagggttagg 180
cgttttgcgc tgcttcgcga tgtacgggcc agatatacgc gttgacattg attattgact 240
agttattaat agtaatcaat tacggggtca ttagttcata gcccatatat ggagttccgc 300
gttacataac ttacggtaaa tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg 360
acgtcaataa tgacgtatgt tcccatagta acgccaatag ggactttcca ttgacgtcaa 420
tgggtggact atttacggta aactgcccac ttggcagtac atcaagtgta tcatatgcca 480
agtacgcccc ctattgacgt caatgacggt aaatggcccg cctggcatta tgcccagtac 540
atgaccttat gggactttcc tacttggcag tacatctacg tattagtcat cgctattacc 600
atggtgatgc ggttttggca gtacaccaat gggcgtggat agcggtttga ctcacgggga 660
tttccaagtc tccaccccat tgacgtcaat gggagtttgt tttggcacca aaatcaacgg 720
gactttccaa aatgtcgtaa caactgcgat cgcccgcccc gttgacgcaa atgggcggta 780
ggcgtgtacg gtgggaggtc tatataagca gagctcgttt agtgaaccgt cagatcacta 840
gaagctttat tgcggtagtt tatcacagtt aaattgctaa cgcagtcagt gcttctgaca 900
caacagtctc gaacttaagc tgcagtgact ctcttaaggt agccttgcag aagttggtcg 960
tgaggcactg ggcaggtaag tatcaaggtt acaagacagg tttaaggaga ccaatagaaa 1020
ctgggcttgt cgagacagag aagactcttg cgtttctgat aggcacctat tggtcttact 1080
gacatccact ttgcctttct ctccacaggt gtccactccc agttcaatta cagctcttaa 1140
ggctagagta cttaatacga ctcactatag gctagcctcg aggtcgacgg tatcgcccga 1200

acagggacct gagaggggcg cagaccctac ctgttgaacc tggctgatcg taggatcccc 1260
gggacagcag aggagaactt acagaagtct tctggaggtg ttcctggcca gaacacagga 1320
ggacaggtaa gatgggagac cctttgacat ggagcaaggc gctcaagaag ttagagaagg 1380
tgacggtaca agggtctcag aaattaacta ctggtaactg taattgggcg ctaagtctag 1440
tagacttatt tcatgatacc aactttgtaa aagaaaagga ctggcagctg agggatgtca 1500
ttccattgct ggaagatgta actcagacgc tgtcaggaca agaaagagag gcctttgaaa 1560
gaacatggtg ggcaatttct gctgtaaaga tgggcctcca gattaataat gtagtagatg 1620
gaaaggcatc attccagctc ctaagagcga aatatgaaaa gaagactgct aataaaaagc 1680
agtctgagcc ctctgaagaa tatccaatca tgatagatgg ggctggaaac agaaatttta 1740
gacctctaac acctagagga tatactactt gggtgaatac catacagaca aatggtctat 1800
taaatgaagc tagtcaaaac ttatttggga tattatcagt agactgtact tctgaagaaa 1860
tgaatgcatt tttggatgtg gtacctggcc aggcaggaca aaagcagata ttacttgatg 1920
caattgataa gatagcagat gattgggata atagacatcc attaccgaat gctccactgg 1980
tggcaccacc acaagggcct attcccatga cagcaaggtt tattagaggt ttaggagtac 2040
ctagagaaag acagatggag cctgcttttg atcagtttag gcagacatat agacaatgga 2100
taatagaagc catgtcagaa ggcatcaaag tgatgattgg aaaacctaaa gctcaaaata 2160
ttaggcaagg agctaaggaa ccttacccag aatttgtaga cagactatta tcccaaataa 2220
aaagtgaggg acatccacaa gagatttcaa aattcttgac tgatacactg actattcaga 2280
acgcaaatga ggaatgtaga aatgctatga gacatttaag accagaggat acattagaag 2340
agaaaatgta tgcttgcaga gacattggaa ctacaaaaca aaagatgatg ttattggcaa 2400
aagcacttca gactggtctt gcgggcccat ttaaaggtgg agccttgaaa ggagggccac 2460
taaaggcagc acaaacatgt tataactgtg ggaagccagg acatttatct agtcaatgta 2520
gagcacctaa agtctgtttt aaatgtaaac agcctggaca tttctcaaag caatgcagaa 2580
gtgttccaaa aaacgggaag caaggggctc aagggaggcc ccagaaacaa actttcccga 2640
tacaacagaa gagtcagcac aacaaatctg ttgtacaaga gactcctcag actcaaaatc 2700
tgtacccaga tctgagcgaa ataaaaaagg aatacaatgt caaggagaag gatcaagtag 2760
aggatctcaa cctggacagt ttgtgggagt aacatataat ctagagaaaa ggcctactac 2820
aatagtatta attaatgata ctcccttaaa tgtactgtta gacacaggag cagatacttc 2880
agtgttgact actgcacatt ataataggtt aaaatataga gggagaaaat atcaagggac 2940
gggaataata ggagtgggag gaaatgtgga aacattttct acgcctgtga ctataaagaa 3000
aaagggtaga cacattaaga caagaatgct agtggcagat attccagtga ctattttggg 3060
acgagatatt cttcaggact taggtgcaaa attggttttg gcacagctct ccaaggaaat 3120
aaaatttaga aaaatagagt taaaagaggg cacaatgggg ccaaaaattc ctcaatggcc 3180
actcactaag gagaaactag aaggggccaa agagatagtc caaagactat tgtcagaggg 3240
aaaaatatca gaagctagtg acaataatcc ttataattca cccatatttg taataaaaaa 3300
gaggtctggc aaatggaggt tattacaaga tctgagagaa ttaaacaaaa cagtacaagt 3360
aggaacggaa atatccagag gattgcctca cccgggagga ttaattaaat gtaaacacat 3420
gactgtatta gatattggag atgcatattt cactataccc ttagatccag agtttagacc 3480
atatacagct ttcactattc cctccattaa tcatcaagaa ccagataaaa gatatgtgtg 3540
gaaatgttta ccacaaggat tcgtgttgag cccatatata tatcagaaaa cattacagga 3600
aattttacaa ccttttaggg aaagatatcc tgaagtacaa ttgtatcaat atatggatga 3660
tttgttcatg ggaagtaatg gttctaaaaa acaacacaaa gagttaatca tagaattaag 3720
ggcgatctta ctggaaaagg gttttgagac accagatgat aaattacaag aagtgccacc 3780
ttatagctgg ctaggttatc aactttgtcc tgaaaattgg aaagtacaaa aaatgcaatt 3840
agacatggta aagaatccaa cccttaatga tgtgcaaaaa ttaatgggga atataacatg 3900
gatgagctca gggatcccag ggttgacagt aaaacacatt gcagctacta ctaagggatg 3960
tttagagttg aatcaaaaag taatttggac ggaagaggca caaaaagagt tagaagaaaa 4020
taatgagaag attaaaaatg ctcaagggtt acaatattat aatccagaag aagaaatgtt 4080
atgtgaggtt gaaattacaa aaaattatga ggcaacttat gttataaaac aatcacaagg 4140
aatcctatgg gcaggtaaaa agattatgaa ggctaataag ggatggtcaa cagtaaaaaa 4200
tttaatgtta ttgttgcaac atgtggcaac agaaagtatt actagagtag gaaaatgtcc 4260
aacgtttaag gtaccattta ccaaagagca agtaatgtgg gaaatgcaaa aaggatggta 4320
ttattcttgg ctcccagaaa tagtatatac acatcaagta gttcatgatg attggagaat 4380
gaaattggta gaagaaccta catcaggaat aacaatatac actgatgggg gaaaacaaaa 4440
tggagaagga atagcagctt atgtgaccag taatgggaga actaaacaga aaaggttagg 4500
acctgtcact catcaagttg ctgaaagaat ggcaatacaa atggcattag aggataccag 4560
agataaacaa gtaaatatag taactgatag ttattattgt tggaaaaata ttacagaagg 4620
attaggttta gaaggaccac aaagtccttg gtggcctata atacaaaata tacgagaaaa 4680

agagatagtt tattttgctt gggtacctgg tcacaaaggg atatatggta atcaattggc 4740
agatgaagcc gcaaaaataa aagaagaaat catgctagca taccaaggca cacaaattaa 4800
agagaaaaga gatgaagatg cagggtttga cttatgtgtt ccttatgaca tcatgatacc 4860
tgtatctgac acaaaaatca tacccacaga tgtaaaaatt caagttcctc ctaatagctt 4920
tggatgggtc actgggaaat catcaatggc aaaacagggg ttattaatta atggaggaat 4980
aattgatgaa ggatatacag gagaaataca agtgatatgt actaatattg gaaaaagtaa 5040
tattaaatta atagagggac aaaaatttgc acaattaatt atactacagc atcactcaaa 5100
ttccagacag ccttgggatg aaaataaaat atctcagaga ggggataaag gatttggaag 5160
tacaggagta ttctgggtag aaaatattca ggaagcacaa gatgaacatg agaattggca 5220
tacatcacca aagatattgg caagaaatta taagatacca ttgactgtag caaaacagat 5280
aactcaagaa tgtcctcatt gcactaagca aggatcagga cctgcaggtt gtgtcatgag 5340
atctcctaat cattggcagg cagattgcac acatttggac aataagataa tattgacttt 5400
tgtagagtca aattcaggat acatacatgc tacattattg tcaaaagaaa atgcattatg 5460
tacttcattg gctattttag aatgggcaag attgttttca ccaaagtcct tacacacaga 5520
taacggcact aattttgtgg cagaaccagt tgtaaatttg ttgaagttcc taaagatagc 5580
acataccaca ggaataccat atcatccaga aagtcagggt attgtagaaa gggcaaatag 5640
gaccttgaaa gagaagattc aaagtcatag agacaacact caaacactgg aggcagcttt 5700
acaacttgct ctcattactt gtaacaaagg gagggaaagt atgggaggac agacaccatg 5760
ggaagtattt atcactaatc aagcacaagt aatacatgag aaacttttac tacagcaagc 5820
acaatcctcc aaaaaatttt gtttttacaa aatccctggt gaacatgatt ggaagggacc 5880
tactagggtg ctgtggaagg gtgatggtgc agtagtagtt aatgatgaag gaaagggaat 5940
aattgctgta ccattaacca ggactaagtt actaataaaa ccaaattgag tattgttgca 6000
ggaagcaaga cccaactacc attgtcagct gtgtttcctg acctcaatat ttgttataag 6060
gtttgatatg aatcccaggg ggaatctcaa cccctattac ccaacagtca gaaaaatcta 6120
agtgtgagga gaacacaatg tttcaacctt attgttataa taatgacagt aagaacagca 6180
tggcagaatc gaaggaagca agagaccaag aaatgaacct gaaagaagaa tctaaagaag 6240
aaaaaagaag aaatgactgg tggaaaatag gtatgtttct gttatgctta gcaggaacta 6300
ctggaggaat actttggtgg tatgaaggac tcccacagca acattatata gggttggtgg 6360
cgataggggg aagattaaac ggatctggcc aatcaaatgc tatagaatgc tggggttcct 6420
tcccggggtg tagaccattt caaaattact tcagttatga gaccaataga agcatgcata 6480
tggataataa tactgctaca ttattagaag ctttaaccaa tataactgct ctataaataa 6540
caaaacagaa ttagaaacat ggaagttagt aaagacttct ggcataactc ctttacctat 6600
ttcttctgaa gctaacactg gactaattag acataagaga gattttggta taagtgcaat 6660
agtggcagct attgtagccg ctactgctat tgctgctagc gctactatgt cttatgttgc 6720
tctaactgag gttaacaaaa taatggaagt acaaaatcat acttttgagg tagaaaatag 6780
tactctaaat ggtatggatt taatagaacg acaaataaag atattatatg ctatgattct 6840
tcaaacacat gcagatgttc aactgttaaa ggaaagacaa caggtagagg agacatttaa 6900
tttaattgga tgtatagaaa gaacacatgt attttgtcat actggtcatc cctggaatat 6960
gtcatgggga catttaaatg agtcaacaca atgggatgac tgggtaagca aaatggaaga 7020
tttaaatcaa gagatactaa ctacacttca tggagccagg aacaatttgg cacaatccat 7080
gataacattc aatacaccag atagtatagc tcaatttgga aaagaccttt ggagtcatat 7140
tggaaattgg attcctggat tgggagcttc cattataaaa tatatagtga tgtttttgct 7200
tatttatttg ttactaacct cttcgcctaa gatcctcagg gccctctgga aggtgaccag 7260
tggtgcaggg tcctccggca gtcgttacct gaagaaaaaa ttccatcaca aacatgcatc 7320
gcgagaagac acctgggacc aggcccaaca caacatacac ctagcaggcg tgaccggtgg 7380
atcaggggac aaatactaca agcagaagta ctccaggaac gactggaatg gagaatcaga 7440
ggagtacaac aggcggccaa agagctgggt gaagtcaatc gaggcatttg gagagagcta 7500
tatttccgag aagaccaaag gggagatttc tcagcctggg gcggctatca acgagcacaa 7560
gaacggctct gggggggaaca atcctcacca agggtcctta gacctggaga ttcgaagcga 7620
aggaggaaac atttatgact gttgcattaa agcccaagaa ggaactctcg ctatcccttg 7680
ctgtggattt ccccttatggc tattttgggg gtcgacccgg gcggccgcac tagaggaatt 7740
cgcccctctc cctccccccc ccctaacgtt actggccgaa gccgcttgga ataaggccgg 7800
tgtgtgtttg tctatatgtg attttccacc atattgccgt cttttggcaa tgtgagggcc 7860
cggaaacctg gccctgtctt cttgacgagc attcctaggg gtctttcccc tctcgccaaa 7920
ggaatgcaag gtctgttgaa tgtcgtgaag gaagcagttc ctctggaagc ttcttgaaga 7980
caaacaacgt ctgtagcgac cctttgcagg cagcggaacc ccccacctgg cgacaggtgc 8040
ctctgcggcc aaaagccacg tgtataagat acacctgcaa aggcggcaca accccagtgc 8100
cacgttgtga gttggatagt tgtggaaaga gtcaaatggc tctcctcaag cgtagtcaac 8160

aaggggctga aggatgccca gaaggtaccc cattgtatgg gaatctgatc tggggcctcg 8220
gtgcacatgc tttacatgtg tttagtcgag gttaaaaaag ctctaggccc cccgaaccac 8280
ggggacgtgg ttttcctttg aaaaacacga tgataagctt gccacaaccc cgtaccaaag 8340
atggatagat ccggaaagcc tgaactcacc gcgacgtctg tcgagaagtt tctgatcgaa 8400
aagttcgaca gcgtctccga cctgatgcag ctctcggagg gcgaagaatc tcgtgctttc 8460
agcttcgatg taggagggcg tggatatgtc ctgcgggtaa atagctgcgc cgatggtttc 8520
tacaaagatc gttatgttta tcggcacttt gcatcggccg cgctcccgat tccggaagtg 8580
cttgacattg gggaattcag cgagagcctg acctattgca tctcccgccg tgcacagggt 8640
gtcacgttgc aagacctgcc tgaaaccgaa ctgcccgctg ttctgcagcc ggtcgcggag 8700
gccatggatg cgatcgctgc ggccgatctt agccagacga gcgggttcgg cccattcgga 8760
ccgcaaggaa tcggtcaata cactacatgg cgtgatttca tatgcgcgat tgctgatccc 8820
catgtgtatc actggcaaac tgtgatggac gacaccgtca gtgcgtccgt cgcgcaggct 8880
ctcgatgagc tgatgctttg ggccgaggac tgccccgaag tccggcacct cgtgcacgcg 8940
gatttcggct ccaacaatgt cctgacggac aatggccgca taacagcggt cattgactgg 9000
agcgaggcga tgttcgggga ttcccaatac gaggtcgcca acatcttctt ctggaggccg 9060
tggttggctt gtatggagca gcagacgcgc tacttcgagc ggaggcatcc ggagcttgca 9120
ggatcgccgc ggctccgggc gtatatgctc cgcattggtc ttgaccaact ctatcagagc 9180
ttggttgacg gcaatttcga tgatgcagct tgggcgcagg gtcgatgcga cgcaatcgtc 9240
cgatccggag ccgggactgt cgggcgtaca caaatcgccc gcagaagcgc ggccgtctgg 9300
accgatggct gtgtagaagt actcgccgat agtggaaacc gacgccccag cactcgtccg 9360
agggcaaagg aatagagtag atgccgaccg aacaagagct gatttcgaga acgcctcagc 9420
cagcaactcg cgcgagccta gcaaggcaaa tgcgagagaa cggccttacg cttggtggca 9480
cagttctcgt ccacagttcg ctaagctcgc tcggctgggt cgcgggaggg ccggtcgcag 9540
tgattcaggc ccttctggat tgtgttggtc cccagggcac gattgtcatg cccacgcact 9600
cgggtgatct gactgatccc gcagattgga gatcgccgcc cgtgcctgcc gattgggtgc 9660
agatctagag ctcgctgatc agcctcgact gtgcctctag ttgccagcca tctgttgttt 9720
gcccctcccc cgtgccttcc ttgaccctgg aaggtgccac tcccactgtc ctttcctaat 9780
aaaatgagga aattgcatcg cattgtctga gtaggtgtca ttctattctg gggggtgggg 9840
tggggcagga cagcaagggg gaggattggg aagacaatag caggcatgct ggggatgcgg 9900
tgggctctat ggcttctgag gcggaaagaa ccagctgggg ctcgagtgca ttctagttgt 9960
ggtttgtcca aactcatcaa tgtatcttat catgtctgta taccgtcgac ctctagctag 10020
agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gctcacaatt 10080
ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc 10140
taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc 10200
cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggcgctct 10260
tccgcttcct cgctcactga ctcgctgcgc tcggtcgttc ggctgcggcg agcggtatca 10320
gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 10380
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 10440
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 10500
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 10560
tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cctttctccc ttcgggaagc 10620
gtggcgcttt ctcaatgctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 10680
aagctgggct gtgtgcacga accccccgtt cagcccgacc gctgcgcctt atccggtaac 10740
tatcgtcttg agtccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 10800
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 10860
aactacggct acactagaag gacagtattt ggtatctgcg ctctgctgaa gccagttacc 10920
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 10980
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 11040
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 11100
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 11160
tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 11220
gcacctatct cagcgatctg tctatttcgt tcatccatag ttgcctgact ccccgtcgtg 11280
tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgcga 11340
gacccacgct caccggctcc agatttatca gcaataaacc agccagccgg aagggccgag 11400
cgcagaagtg gtcctgcaac tttatccgcc tccatccagt ctattaattg ttgccgggaa 11460
gctagagtaa gtagttcgcc agttaatagt ttgcgcaacg ttgttgccat tgctacaggc 11520
atcgtggtgt cacgctcgtc gtttggtatg gcttcattca gctccggttc ccaacgatca 11580
aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggtcctccg 11640

atcgttgtca gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 11700
aattctctta ctgtcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 11760
aagtcattct gagaatagtg tatgcggcga ccgagttgct cttgcccggc gtcaatacgg 11820
gataataccg cgccacatag cagaacttta aaagtgctca tcattggaaa acgttcttcg 11880
gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 11940
gcacccaact gatcttcagc atcttttact ttcaccagcg tttctgggtg agcaaaaaca 12000
ggaaggcaaa atgccgcaaa aaagggaata agggcgacac ggaaatgttg aatactcata 12060
ctcttccttt ttcaatatta ttgaagcatt tatcagggtt attgtctcat gagcggatac 12120
atatttgaat gtatttagaa aaataaacaa ataggggttc cgcgcacatt tccccgaaaa 12180
gtgccacctg acgtcgacgg atcggg 12206

<210> 52
<211> 11592
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 52
agatcttgaa taataaaatg tgtgtttgtc cgaaatacgc gttttgagat ttctgtcgcc 60
gactaaattc atgtcgcgcg atagtggtgt ttatcgccga tagagatggc gatattggaa 120
aaattgatat ttgaaaatat ggcatattga aaatgtcgcc gatgtgagtt tctgtgtaac 180
tgatatcgcc atttttccaa aagtgatttt tgggcatacg cgatatctgg cgatagcgct 240
tatatcgttt acgggggatg gcgatagacg actttggtga cttgggcgat tctgtgtgtc 300
gcaaatatcg cagtttcgat ataggtgaca gacgatatga ggctatatcg ccgatagagg 360
cgacatcaag ctggcacatg gccaatgcat atcgatctat acattgaatc aatattggcc 420
attagccata ttattcattg gttatatagc ataaatcaat attggctatt ggccattgca 480
tacgttgtat ccatatcgta atatgtacat ttatattggc tcatgtccaa cattaccgcc 540
atgttgacat tgattattga ctagttatta atagtaatca attacggggt cattagttca 600
tagcccatat atggagttcc gcgttacata acttacggta aatggcccgc ctggctgacc 660
gcccaacgac ccccgcccat tgacgtcaat aatgacgtat gttcccatag taacgccaat 720
agggactttc cattgacgtc aatgggtgga gtatttacgg taaactgccc acttggcagt 780
acatcaagtg tatcatatgc caagtccgcc ccctattgac gtcaatgacg gtaaatggcc 840
cgcctggcat tatgcccagt acatgacctt acgggacttt cctacttggc agtacatcta 900
cgtattagtc atcgctatta ccatggtgat gcggttttgg cagtacacca atgggcgtgg 960
atagcggttt gactcacggg gatttccaag tctccacccc attgacgtca atgggagttt 1020
gttttggcac caaaatcaac gggactttcc aaaatgtcgt aacaactgcg atcgcccgcc 1080
ccgttgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag cagagctcgt 1140
ttagtgaacc gggcactcag attctgcggt ctgagtccct tctctgctgg gctgaaaagg 1200
cctttgtaat aaatataatt ctctactcag tccctgtctc tagtttgtct gttcgagatc 1260
ctacagttgg cgcccgaaca gggacctgag aggggcgcag accctacctg ttgaacctgg 1320
ctgatcgtag gatccccggg acagcagagg agaacttaca gaagtcttct ggaggtgttc 1380
ctggccagaa cacaggagga caggtaagat tgggagaccc tttgacattg gagcaaggcg 1440
ctcaagaagt tagagaaggt gacggtacaa gggtctcaga aattaactac tggtaactgt 1500
aattgggcgc taagtctagt agacttattt cattgatacc aactttgtaa aagaaaagga 1560
ctggcagctg agggattgtc attccattgc tggaagattg taactcagac gctgtcagga 1620
caagaaagag aggcctttga aagaacattg gtgggcaatt tctgctgtaa agattgggcc 1680
tccagattaa taattgtagt agattggaaa ggcatcattc cagctcctaa gagcgaaata 1740
ttgaaaagaa gactgctaat aaaaagcagt ctgagccctc tgaagaatat ctctagaact 1800
agtggatccc ccgggctgca ggagtgggga ggcacgatgg ccgctttggt cgaggcggat 1860
ccggccatta gccatattat tcattggtta tatagcataa atcaatattg gctattggcc 1920
attgcatacg ttgtatccat atcataatat gtacatttat attggctcat gtccaacatt 1980
accgccatgt tgacattgat tattgactag ttattaatag taatcaatta cggggtcatt 2040
agttcatagc ccatatatgg agttccgcgt tacataactt acggtaaatg gcccgcctgg 2100
ctgaccgccc aacgaccccc gcccattgac gtcaataatg acgtatgttc ccatagtaac 2160

gccaataggg actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt 2220
ggcagtacat caagtgtatc atatgccaag tacgccccct attgacgtca atgacggtaa 2280
atggcccgcc tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta 2340
catctacgta ttagtcatcg ctattaccat ggtgatgcgg ttttggcagt acatcaatgg 2400
gcgtggatag cggtttgact cacggggatt tccaagtctc caccccattg acgtcaatgg 2460
gagtttgttt tggcaccaaa atcaacggga ctttccaaaa tgtcgtaaca actccgcccc 2520
attgacgcaa atgggcggta ggcatgtacg gtgggaggtc tatataagca gagctcgttt 2580
agtgaaccgt cagatcgcct ggagacgcca tccacgctgt tttgacctcc atagaagaca 2640
ccgggaccga tccagcctcc gcggccccaa gcttcagctg ctcgaggatc tgcggatccg 2700
gggaattccc cagtctcagg atccaccatg gggatcccg tcgttttaca acgtcgtgac 2760
tgggaaaacc ctggcgttac ccaacttaat cgccttgcag cacatccccc tttcgccagc 2820
tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc aacagttgcg cagcctgaat 2880
ggcgaatggc gctttgcctg gtttccggca ccagaagcgg tgccggaaag ctggctggag 2940
tgcgatcttc ctgaggccga tactgtcgtc gtcccctcaa actggcagat gcacggttac 3000
gatgcgccca tctacaccaa cgtaacctat cccattacgg tcaatccgcc gtttgttccc 3060
acggagaatc cgacgggttg ttactcgctc acatttaatg ttgatgaaag ctggctacag 3120
gaaggccaga cgcgaattat ttttgatggc gttaactcgg cgtttcatct gtggtgcaac 3180
gggcgctggg tcggttacgg ccaggacagt cgtttgccgt ctgaatttga cctgagcgca 3240
tttttacgcg ccggagaaaa ccgcctcgcg gtgatggtgc tgcgttggag tgacggcagt 3300
tatctggaag atcaggatat gtggcggatg agcggcattt tccgtgacgt ctcgttgctg 3360
cataaaccga ctacacaaat cagcgatttc catgttgcca ctcgctttaa tgatgatttc 3420
agccgcgctg tactggaggc tgaagttcag atgtgcggcg agttgcgtga ctacctacgg 3480
gtaacagttt ctttatggca gggtgaaacg caggtcgcca gcggcaccgc gcctttcggc 3540
ggtgaaatta tcgatgagcg tggtggttat gccgatcgcg tcacactacg tctgaacgtc 3600
gaaaacccga aactgtggag cgccgaaatc ccgaatctct atcgtgcggt ggttgaactg 3660
cacaccgccg acggcacgct gattgaagca gaagcctgcg atgtcggttt ccgcgaggtg 3720
cggattgaaa atggtctgct gctgctgaac ggcaagccgt tgctgattcg aggcgttaac 3780
cgtcacgagc atcatcctct gcatggtcag gtcatggatg agcagacgat ggtgcaggat 3840
atcctgctga tgaagcagaa caactttaac gccgtgcgct gttcgcatta tccgaaccat 3900
ccgctgtggt acacgctgtg cgaccgctac ggcctgtatg tggtggatga agccaatatt 3960
gaaacccacg gcatggtgcc aatgaatcgt ctgaccgatg atccgcgctg gctaccggcg 4020
atgagcgaac gcgtaacgcg aatggtgcag cgcgatcgta atcacccgag tgtgatcatc 4080
tggtcgctgg ggaatgaatc aggccacggc gctaatcacg acgcgctgta tcgctggatc 4140
aaatctgtcg atccttcccg cccggtgcag tatgaaggcg gcggagccga caccacggcc 4200
accgatatta tttgcccgat gtacgcgcgc gtggatgaag accagccctt cccggctgtg 4260
ccgaaatggt ccatcaaaaa atggctttcg ctacctggag agacgcgccc gctgatcctt 4320
tgcgaatacg cccacgcgat gggtaacagt cttggcggtt tcgctaaata ctggcaggcg 4380
tttcgtcagt atccccgttt acagggcggc ttcgtctggg actgggtgga tcagtcgctg 4440
attaaatatg atgaaaacgg caacccgtgg tcggcttacg gcggtgattt tggcgatacg 4500
ccgaacgatc gccagttctg tatgaacggt ctggtctttg ccgaccgcac gccgcatcca 4560
gcgctgacgg aagcaaaaca ccagcagcag tttttccagt tccgtttatc cgggcaaacc 4620
atcgaagtga ccagcgaata cctgttccgt catagcgata acgagctcct gcactggatg 4680
gtggcgctgg atggtaagcc gctggcaagc ggtgaagtgc ctctggatgt cgctccacaa 4740
ggtaaacagt tgattgaact gcctgaacta ccgcagccgg agagcgccgg gcaactctgg 4800
ctcacagtac gcgtagtgca accgaacgcg accgcatggt cagaagccgg gcacatcagc 4860
gcctggcagc agtggcgtct ggcggaaaac ctcagtgtga cgctccccgc cgcgtcccac 4920
gccatcccgc atctgaccac cagcgaaatg gatttttgca tcgagctggg taataagcgt 4980
tggcaattta accgccagtc aggctttctt tcacagatgt ggattggcga taaaaaacaa 5040
ctgctgacgc cgctgcgcga tcagttcacc cgtgcaccgc tggataacga cattggcgta 5100
agtgaagcga cccgcattga ccctaacgcc tgggtcgaac gctggaaggc ggcgggccat 5160
taccaggccg aagcagcgtt gttgcagtgc acggcagata cacttgctga tgcggtgctg 5220
attacgaccg ctcacgcgtg gcagcatcag gggaaaacct tatttatcag ccggaaaacc 5280
taccggattg atggtagtgg tcaaatggcg attaccgttg atgttgaagt ggcgagcgat 5340
acaccgcatc cggcgcggat tggcctgaac tgccagctgg cgcaggtagc agagcgggta 5400
aactggctcg gattagggcc gcaagaaaac tatcccgacc gccttactgc cgcctgtttt 5460
gaccgctggg atctgccatt gtcagacatg tataccccgt acgtcttccc gagcgaaaac 5520
ggtctgcgct gcgggacgcg cgaattgaat tatggcccac accagtggcg cggcgacttc 5580
cagttcaaca tcagccgcta cagtcaacag caactgatgg aaaccagcca tcgccatctg 5640

```
ctgcacgcgg aagaaggcac atggctgaat atcgacggtt tccatatggg gattggtggc   5700
gacgactcct ggagcccgtc agtatcggcg gaattccagc tgagcgccgg tcgctaccat   5760
taccagttgg tctggtgtca aaaataataa taaccgggca ggggggatcc gcagatccgg   5820
ctgtggaatg tgtgtcagtt agggtgtgga aagtccccag gctccccagc aggcagaagt   5880
atgcaaagca tgcctgcagg aattcgatat caagcttatc gataccgtcg acctcgaggg   5940
ggggcccggt acccagcttt tgttcccttt agtgagggtt aattgcgcgg gaagtattta   6000
tcactaatca agcacaagta atacatgaga aacttttact acagcaagca caatcctcca   6060
aaaaattttg tttttacaaa atccctggtg aacatgattg gaagggacct actagggtgc   6120
tgtggaaggg tgatggtgca gtagtagtta atgatgaagg aaagggaata attgctgtac   6180
cattaaccag gactaagtta ctaataaaac caaattgagt attgttgcag gaagcaagac   6240
ccaactacca ttgtcagctg tgtttcctga cctcaatatt tgttataagg tttgatatga   6300
atcccagggg gaatctcaac ccctattacc caacagtcag aaaaatctaa gtgtgaggag   6360
aacacaatgt ttcaacctta ttgttataat aatgacagta agaacagcat ggcagaatcg   6420
aaggaagcaa gagaccaaga atgaacctga aagaagaatc taaagaagaa aaaagaagaa   6480
atgactggtg gaaaataggt atgtttctgt tatgcttagc aggaactact ggaggaatac   6540
tttggtggta tgaaggactc ccacagcaac attatatagg gttggtggcg ataggggggaa  6600
gattaaacgg atctggccaa tcaaatgcta tagaatgctg ggttccttc ccggggtgta   6660
gaccatttca aaattacttc agttatgaga ccaatagaag catgcatatg gataataata   6720
ctgctacatt attagaagct ttaaccaata taactgctct ataaataaca aaacagaatt   6780
agaaacatgg aagttagtaa agacttctgg cataactcct ttacctattt cttctgaagc   6840
taacactgga ctaattagac ataagagaga ttttggtata agtgcaatag tggcagctat   6900
tgtagccgct actgctattg ctgctagcgc tactatgtct tatgttgctc taactgaggt   6960
taacaaaata atggaagtac aaaatcatac ttttgaggta gaaaatagta ctctaaatgg   7020
tatggattta atagaacgac aaataaagat attatatgct atgattcttc aaacacatgc   7080
agatgttcaa ctgttaaagg aaagacaaca ggtagaggag acatttaatt taattggatg   7140
tatagaaaga acacatgtat tttgtcatac tggtcatccc tggaatatgt catggggaca   7200
tttaaatgag tcaacacaat gggatgactg ggtaagcaaa atggaagatt taaatcaaga   7260
gatactaact acacttcatg gagccaggaa caatttggca caatccatga taacattcaa   7320
tacaccagat agtatagctc aatttggaaa agacctttgg agtcatattg gaaattggat   7380
tcctggattg ggagcttcca ttataaaata tatagtgatg tttttgctta tttatttgtt   7440
actaacctct tcgcctaaga tcctcagggc cctctggaag gtgaccagtg gtgcagggtc   7500
ctccggcagt cgttacctga agaaaaaatt ccatcacaaa catgcatcgc gagaagacac   7560
ctgggaccag gcccaacaca acatacacct agcaggcgtg accggtggat caggggacaa   7620
atactacaag cagaagtact ccaggaacga ctggaatgga gaatcagagg agtacaacag   7680
gcggccaaag agctgggtga agtcaatcga ggcatttgga gagagctata tttccgagaa   7740
gaccaaaggg gagatttctc agcctggggc ggctatcaac gagcacaaga acggctctgg   7800
ggggaacaat cctcaccaag ggtccttaga cctggagatt cgaagcgaag gaggaaacat   7860
ttatgactgt tgcattaaag cccaagaagg aactctcgct atcccttgct gtggatttcc   7920
cttatggcta ttttggggac tagtaattat agtaggacgc atagcaggct atggattacg   7980
tggactcgct gttataataa ggatttgtat tagaggctta aatttgatat ttgaaataat   8040
cagaaaaatg cttgattata ttggaagagc tttaaatcct ggcacatctc atgtatcaat   8100
gcctcagtat gtttagaaaa acaagggggg aactgtgggg tttttatgag ggttttata    8160
aactgcagga gtggggaggc acgatggccg ctttggtcga ggcggatccg gccattagcc   8220
atattattca ttggttatat agcataaatc aatattggct attggccatt gcatacgttg   8280
tatccatatc ataatatgta catttatatt ggctcatgtc caacattacc gccatgttga   8340
cattgattat tgactagtta ttaatagtaa tcaattacgg ggtcattagt tcatagccca   8400
tatatggagt tccgcgttac ataacttacg gtaaatggcc cgcctggctg accgcccaac   8460
gacccccgcc cattgacgtc aataatgacg tatgttccca tagtaacgcc aatagggact   8520
ttccattgac gtcaatgggt ggagtattta cggtaaactg cccacttggc agtacatcaa   8580
gtgtatcata tgccaagtac gccccctatt gacgtcaatg acggtaaatg gcccgcctgg   8640
cattatgccc agtacatgac cttatgggac tttcctactt ggcagtacat ctacgtatta   8700
gtcatcgcta ttaccatggt gatgcggttt tggcagtaca tcaatgggcg tggatagcgg   8760
tttgactcac ggggatttcc aagtctccac cccattgacg tcaatgggag tttgttttgg   8820
caccaaaatc aacgggactt tccaaaatgt cgtaacaact ccgccccatt gacgcaaatg   8880
ggcggtaggc atgtacggtg ggaggtctat ataagcagag ctcgtttagt gaaccgggca   8940
ctcagattct gcggtctgag tcccttctct gctgggctga aaggcctttt gtaataaata   9000
taattctcta ctcagtccct gtctctagtt tgtctgttcg agatcctaca gagctcatgc   9060
cttggcgtaa tcatggtcat agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc   9120
```

acacaacata cgagccggaa gcataaagtg taaagcctgg ggtgcctaat gagtgagcta 9180
actcacatta attgcgttgc gctcactgcc cgctttccag tcgggaaacc tgtcgtgcca 9240
gctgcattaa tgaatcggcc aacgcgcggg gagaggcggt ttgcgtattg ggcgctcttc 9300
cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg ctgcggcgag cggtatcagc 9360
tcactcaaag gcggtaatac ggttatccac agaatcaggg gataacgcag gaaagaacat 9420
gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag gccgcgttgc tggcgttttt 9480
ccataggctc cgcccccctg acgagcatca caaaaatcga cgctcaagtc agaggtggcg 9540
aaacccgaca ggactataaa gataccaggc gtttccccct ggaagctccc tcgtgcgctc 9600
tcctgttccg accctgccgc ttaccggata cctgtccgcc tttctccctt cgggaagcgt 9660
ggcgctttct catagctcac gctgtaggta tctcagttcg gtgtaggtcg ttcgctccaa 9720
gctgggctgt gtgcacgaac cccccgttca gcccgaccgc tgcgccttat ccggtaacta 9780
tcgtcttgag tccaacccgg taagacacga cttatcgcca ctggcagcag ccactggtaa 9840
caggattagc agagcgaggt atgtaggcgg tgctacagag ttcttgaagt ggtggcctaa 9900
ctacggctac actagaagga cagtatttgg tatctgcgct ctgctgaagc cagttacctt 9960
cggaaaaaga gttggtagct cttgatccgg caaacaaacc accgctggta gcggtggttt 10020
ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga tctcaagaag atcctttgat 10080
cttttctacg gggtctgacg ctcagtggaa cgaaaactca cgttaaggga ttttggtcat 10140
gagattatca aaaaggatct tcacctagat ccttttaaat taaaaatgaa gttttaaatc 10200
aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa tcagtgaggc 10260
acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc ccgtcgtgta 10320
gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga taccgcgaga 10380
cccacgctca ccggctccag atttatcagc aataaaccag ccagccggaa gggccgagcg 10440
cagaagtggt cctgcaactt tatccgcctc catccagtct attaattgtt gccgggaagc 10500
tagagtaagt agttcgccag ttaatagttt gcgcaacgtt gttgccattg ctacaggcat 10560
cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc tccggttccc aacgatcaag 10620
gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt agctccttcg gtcctccgat 10680
cgttgtcaga agtaagttgg ccgcagtgtt atcactcatg gttatggcag cactgcataa 10740
ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt actcaaccaa 10800
gtcattctga gaatagtgta tgcggcgacc gagttgctct tgcccggcgt caatacggga 10860
taataccgcg ccacatagca gaactttaaa agtgctcatc attggaaaac gttcttcggg 10920
gcgaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac ccactcgtgc 10980
acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag caaaaacagg 11040
aaggcaaaat gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatact 11100
cttccttttt caatattatt gaagcattta tcagggttat tgtctcatga gcggatacat 11160
atttgaatgt atttagaaaa ataaacaaat aggggttccg cgcacatttc cccgaaaagt 11220
gccacctaaa ttgtaagcgt taatattttg ttaaaattcg cgttaaattt ttgttaaatc 11280
agctcatttt ttaaccaata ggccgaaatc ggcaaaatcc cttataaatc aaaagaatag 11340
accgagatag ggttgagtgt tgttccagtt tggaacaaga gtccactatt aaagaacgtg 11400
gactccaacg tcaaagggcg aaaaaccgtc tatcagggcg atggcccact acgtgaacca 11460
tcaccctaat caagtttttt ggggtcgagg tgccgtaaag cactaaatcg gaaccctaaa 11520
gggagccccc gatttagagc ttgacgggga aagccaacct ggcttatcga aattaatacg 11580
actcactata gg 11592

<210> 53
<211> 10112
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 53
tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60
ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc 120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg 180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc 240

gcctggctga ccgcccaacg acccccgccc attgacgtca ataatgacgt atgttcccat 300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc 360
ccacttggca gtacatcaag tgtatcatat gccaagtccg ccccctattg acgtcaatga 420
cggtaaatgg cccgcctggc attatgccca gtacatgacc ttacgggact ttcctacttg 480
gcagtacatc tacgtattag tcatcgctat taccatggtg atgcggtttt ggcagtacac 540
caatgggcgt ggatagcggt ttgactcacg gggatttcca agtctccacc ccattgacgt 600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactg 660
cgatcgcccg ccccgttgac gcaaatgggc ggtaggcgtg tacggtggga ggtctatata 720
agcagagctc gtttagtgaa ccgtcagatc actagaagct ttattgcggt agtttatcac 780
agttaaattg ctaacgcagt cagtgcttct gacacaacag tctcgaactt aagctgcagt 840
gactctctta aggtagcctt gcagaagttg gtcgtgaggc actgggcagg taagtatcaa 900
ggttacaaga caggtttaag gagaccaata gaaactgggc ttgtcgagac agagaagact 960
cttgcgtttc tgataggcac ctattggtct tactgacatc cactttgcct ttctctccac 1020
aggtgtccac tcccagttca attacagctc ttaaggctag agtacttaat acgactcact 1080
ataggctaga gaattcgcca ccatgggcga tcccctcacc tggtccaaag ccctgaagaa 1140
actggaaaaa gtcaccgttc agggtagcca aaagcttacc acaggcaatt gcaactgggc 1200
attgtccctg gtggatcttt tccacgacac taatttcgtt aaggagaaag attggcaact 1260
cagagacgtg atcccccctct tggaggacgt gacccaaaca ttgtctgggc aggagcgcga 1320
agctttcgag cgcacctggt gggccatcag cgcagtcaaa atggggctgc aaatcaacaa 1380
cgtggttgac ggtaaagcta gctttcaact gctccgcgct aagtacgaga agaaaaccgc 1440
caacaagaaa caatccgaac ctagcgagga gtacccaatt atgatcgacg gcgccggcaa 1500
taggaacttc cgcccactga ctcccagggg ctataccacc tgggtcaaca ccatccagac 1560
aaacggactt ttgaacgaag cctcccagaa cctgttcggc atcctgtctg tggactgcac 1620
ctccgaagaa atgaatgctt ttctcgacgt ggtgccagga caggctggac agaaacagat 1680
cctgctcgat gccattgaca agatcgccga cgactgggat aatcgccacc ccctgccaaa 1740
cgcccctctg gtggctcccc cacaggggcc tatccctatg accgctaggt tcattagggg 1800
actgggggtg ccccgcgaac gccagatgga gccagcattt gaccaattta ggcagaccta 1860
cagacagtgg atcatcgaag ccatgagcga ggggattaaa gtcatgatcg gaaagcccaa 1920
ggcacagaac atcaggcagg gggccaagga accatacccct gagtttgtcg acaggcttct 1980
gtcccagatt aaatccgaag gccaccctca ggagatctcc aagttcttga cagacacact 2040
gactatccaa aatgcaaatg aagagtgcag aaacgccatg aggcacctca gacctgaaga 2100
taccctggag gagaaaatgt acgcatgtcg cgacattggc actaccaagc aaaagatgat 2160
gctgctcgcc aaggctctgc aaaccggcct ggctggtcca ttcaaaggag gagcactgaa 2220
gggaggtcca ttgaaagctg cacaaacatg ttataattgt gggaagccag gacatttatc 2280
tagtcaatgt agagcaccta aagtctgttt taaatgtaaa cagcctggac atttctcaaa 2340
gcaatgcaga agtgttccaa aaaacgggaa gcaaggggct caagggaggc cccagaaaca 2400
aactttcccg atacaacaga agagtcagca caacaaatct gttgtacaag agactcctca 2460
gactcaaaat ctgtacccag atctgagcga aataaaaaag gaatacaatg tcaaggagaa 2520
ggatcaagta gaggatctca acctggacag tttgtgggag taacatacaa tctcgagaag 2580
aggcccacta ccatcgtcct gatcaatgac acccctctta atgtgctgct ggacaccgga 2640
gccgacacca gcgttctcac tactgctcac tataacagac tgaaatacag aggaaggaaa 2700
taccagggca caggcatcat cggcgttgga ggcaacgtcg aaacctttc cactcctgtc 2760
accatcaaaa agaaggggag acacattaaa accagaatgc tggtcgccga catccccgtc 2820
accatccttg gcagagacat tctccaggac ctgggcgcta aactcgtgct ggcacaactg 2880
tctaaggaaa tcaagttccg caagatcgag ctgaaagagg gcacaatggg tccaaaaatc 2940
ccccagtggc cctgaccaa agagaagctt gagggcgcta aggaaatcgt gcagcgcctg 3000
ctttctgagg gcaagattag cgaggccagc gacaataacc cttacaacag ccccatcttt 3060
gtgattaaga aaaggagcgg caaatggaga ctcctgcagg acctgaggga actcaacaag 3120
accgtccagg tcggaactga gatctctcgc ggactgcctc accccggcgg cctgattaaa 3180
tgcaagcaca tgacagtcct tgacattgga gacgcttatt ttaccatccc cctcgatcct 3240
gaatttcgcc cctatactgc ttttaccatc cccagcatca atcaccagga gcccgataaa 3300
cgctatgtgt ggaagtgcct cccccaggga tttgtgctta gcccctacat ttaccagaag 3360
acacttcaag agatcctcca accttttccgc gaaagatacc cagaggttca actctaccaa 3420
tatatggacg acctgttcat ggggtccaac gggtctaaga agcagcacaa ggaactcatc 3480
atcgaactga gggcaatcct cctggagaaa ggcttcgaga cacccgacga caagctgcaa 3540
gaagttcctc catatagctg gctgggctac cagctttgcc ctgaaaactg gaaagtccag 3600
aagatgcagt tggatatggt caagaaccca acactgaacg acgtccagaa gctcatgggc 3660
aatattacct ggatgagctc cggaatccct gggcttaccg ttaagcacat tgccgcaact 3720

acaaaaggat gcctggagtt gaaccagaag gtcatttgga cagaggaagc tcagaaggaa 3780
ctggaggaga ataatgaaaa gattaagaat gctcaagggc tccaatacta caatcccgaa 3840
gaagaaatgt tgtgcgaggt cgaaatcact aagaactacg aagccaccta tgtcatcaaa 3900
cagtcccaag gcatcttgtg ggccggaaag aaaatcatga aggccaacaa aggctggtcc 3960
accgttaaaa atctgatgct cctgctccag cacgtcgcca ccgagtctat cacccgcgtc 4020
ggcaagtgcc ccaccttcaa agttcccttc actaaggagc aggtgatgtg ggagatgcaa 4080
aaaggctggt actactcttg gcttcccgag atcgtctaca cccaccaagt ggtgcacgac 4140
gactggagaa tgaagcttgt cgaggagccc actagcggaa ttacaatcta taccgacggc 4200
ggaaagcaaa acggagaggg aatcgctgca tacgtcacat ctaacggccg caccaagcaa 4260
aagaggctcg gccctgtcac tcaccaggtg gctgagagga tggctatcca gatggccctt 4320
gaggacacta gagacaagca ggtgaacatt gtgactgaca gctactactg ctggaaaaac 4380
atcacagagg gccttggcct ggagggaccc cagtctccct ggtggcctat catccagaat 4440
atccgcgaaa aggaaattgt ctatttcgcc tgggtgcctg gacacaaagg aatttacggc 4500
aaccaactcg ccgatgaagc cgccaaaatt aaagaggaaa tcatgcttgc ctaccagggc 4560
acacagatta aggagaagag agacgaggac gctggctttg acctgtgtgt gccatacgac 4620
atcatgattc ccgttagcga cacaaagatc attccaaccg atgtcaagat ccaggtgcca 4680
cccaattcat ttggttgggt gaccggaaag tccagcatgg ctaagcaggg tcttctgatt 4740
aacgggggaa tcattgatga aggatacacc ggcgaaatcc aggtgatctg cacaaatatc 4800
ggcaaaagca atattaagct tatcgaaggg cagaagttcg ctcaactcat catcctccag 4860
caccacagca attcaagaca accttgggac gaaaacaaga ttagccagag aggtgacaag 4920
ggcttcggca gcacaggtgt gttctgggtg gagaacatcc aggaagcaca ggacgagcac 4980
gagaattggc acacctcccc taagattttg gcccgcaatt acaagatccc actgactgtg 5040
gctaagcaga tcacacagga atgcccccac tgcaccaaac aaggttctgg ccccgccggc 5100
tgcgtgatga ggtcccccaa tcactggcag gcagattgca cccacctcga caacaaaatt 5160
atcctgacct tcgtggagag caattccggc tacatccacg caacactcct ctccaaggaa 5220
aatgcattgt gcacctccct cgcaattctg gaatgggcca ggctgttctc tccaaaatcc 5280
ctgcacaccg acaacggcac caactttgtg gctgaacctg tggtgaatct gctgaagttc 5340
ctgaaaatcg cccacaccac tggcattccc tatcaccctg aaagccaggg cattgtcgag 5400
agggccaaca gaactctgaa agaaaagatc caatctcaca gagacaatac acagacattg 5460
gaggccgcac ttcagctcgc ccttatcacc tgcaacaaag gaagagaaag catggcggc 5520
cagacccccct gggaggtctt catcactaac caggcccagg tcatccatga aaagctgctc 5580
ttgcagcagg cccagtcctc caaaaagttc tgcttttata agatccccgg tgagcacgac 5640
tggaaaggtc ctacaagagt tttgtggaaa ggagacggcg cagttgtggt gaacgatgag 5700
ggcaagggga tcatcgctgt gcccctgaca cgcaccaagc ttctcatcaa gccaaactga 5760
acccggggcg gccgcttccc tttagtgagg gttaatgctt cgagcagaca tgataagata 5820
cattgatgag tttggacaaa ccacaactag aatgcagtga aaaaaatgct ttatttgtga 5880
aatttgtgat gctattgctt tatttgtaac cattataagc tgcaataaac aagttaacaa 5940
caacaattgc attcatttta tgtttcaggt tcagggggag atgtgggagg ttttttaaag 6000
caagtaaaac ctctacaaat gtggtaaaat ccgataagga tcgatccggg ctggcgtaat 6060
agcgaagagg cccgcaccga tcgcccttcc caacagttgc gcagcctgaa tggcgaatgg 6120
acgcgccctg tagcggcgca ttaagcgcgg cgggtgtggt ggttacgcgc agcgtgaccg 6180
ctacacttgc cagcgcccta gcgcccgctc ctttcgcttt cttcccttcc tttctcgcca 6240
cgttcgccgg ctttccccgt caagctctaa atcgggggct ccctttaggg ttccgattta 6300
gagctttacg gcacctcgac cgcaaaaaac ttgatttggg tgatggttca cgtagtgggc 6360
catcgccctg atagacggtt tttcgccctt tgacgttgga gtccacgttc tttaatagtg 6420
gactcttgtt ccaaactgga acaacactca accctatctc ggtctattct tttgatttat 6480
aagggatttt gccgatttcg gcctattggt taaaaaatga gctgatttaa caaatattta 6540
acgcgaattt taacaaaata ttaacgttta caatttcgcc tgatgcggta ttttctcctt 6600
acgcatctgt gcggtatttc acaccgcata cgcggatctg cgcagcacca tggcctgaaa 6660
taacctctga aagaggaact tggttaggta ccttctgagg cggaaagaac cagctgtgga 6720
atgtgtgtca gttagggtgt ggaaagtccc caggctcccc agcaggcaga agtatgcaaa 6780
gcatgcatct caattagtca gcaaccaggt gtggaaagtc cccaggctcc ccagcaggca 6840
gaagtatgca aagcatgcat ctcaattagt cagcaaccat agtcccgccc ctaactccgc 6900
ccatcccgcc cctaactccg cccagttccg cccattctcc gccccatggc tgactaattt 6960
tttttattta tgcagaggcc gaggccgcct cggcctctga gctattccag aagtagtgag 7020
gaggcttttt tggaggccta ggcttttgca aaaagcttga ttcttctgac acaacagtct 7080
cgaacttaag gctagagcca ccatgattga acaagatgga ttgcacgcag gttctccggc 7140
cgcttgggtg gagaggctat tcggctatga ctgggcacaa cagacaatcg gctgctctga 7200

tgccgccgtg ttccggctgt cagcgcaggg gcgcccggtt ctttttgtca agaccgacct 7260
gtccggtgcc ctgaatgaac tgcaggacga ggcagcgcgg ctatcgtggc tggccacgac 7320
gggcgttcct tgcgcagctg tgctcgacgt tgtcactgaa gcgggaaggg actggctgct 7380
attgggcgaa gtgccggggc aggatctcct gtcatctcac cttgctcctg ccgagaaagt 7440
atccatcatg gctgatgcaa tgcggcggct gcatacgctt gatccggcta cctgcccatt 7500
cgaccaccaa gcgaaacatc gcatcgagcg agcacgtact cggatggaag ccggtcttgt 7560
cgatcaggat gatctggacg aagagcatca ggggctcgcg ccagccgaac tgttcgccag 7620
gctcaaggcg cgcatgcccg acggcgagga tctcgtcgtg acccatggcg atgcctgctt 7680
gccgaatatc atggtggaaa atggccgctt ttctggattc atcgactgtg gccggctggg 7740
tgtggcggac cgctatcagg acatagcgtt ggctacccgt gatattgctg aagagcttgg 7800
cggcgaatgg gctgaccgct tcctcgtgct ttacggtatc gccgctcccg attcgcagcg 7860
catcgccttc tatcgccttc ttgacgagtt cttctgagcg ggactctggg gttcgaaatg 7920
accgaccaag cgacgcccaa cctgccatca cgatggccgc aataaaatat ctttattttc 7980
attacatctg tgtgttggtt ttttgtgtga atcgatagcg ataaggatcc gcgtatggtg 8040
cactctcagt acaatctgct ctgatgccgc atagttaagc cagccccgac acccgccaac 8100
acccgctgac gcgccctgac gggcttgtct gctcccggca tccgcttaca gacaagctgt 8160
gaccgtctcc gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgag 8220
acgaaagggc ctcgtgatac gcctattttt ataggttaat gtcatgataa taatggtttc 8280
ttagacgtca ggtggcactt ttcggggaaa tgtgcgcgga acccctattt gtttattttt 8340
ctaaatacat tcaaatatgt atccgctcat gagacaataa ccctgataaa tgcttcaata 8400
atattgaaaa aggaagagta tgagtattca acatttccgt gtcgccctta ttcccttttt 8460
tgcggcattt tgccttcctg tttttgctca cccagaaacg ctggtgaaag taaaagatgc 8520
tgaagatcag ttgggtgcac gagtgggtta catcgaactg gatctcaaca gcggtaagat 8580
ccttgagagt tttcgccccg aagaacgttt tccaatgatg agcactttta aagttctgct 8640
atgtggcgcg gtattatccc gtattgacgc cgggcaagag caactcggtc gccgcataca 8700
ctattctcag aatgacttgg ttgagtactc accagtcaca gaaaagcatc ttacggatgg 8760
catgacagta agagaattat gcagtgctgc cataaccatg agtgataaca ctgcggccaa 8820
cttacttctg acaacgatcg gaggaccgaa ggagctaacc gcttttttgc acaacatggg 8880
ggatcatgta actcgccttg atcgttggga accggagctg aatgaagcca taccaaacga 8940
cgagcgtgac accacgatgc ctgtagcaat ggcaacaacg ttgcgcaaac tattaactgg 9000
cgaactactt actctagctt cccggcaaca attaatagac tggatggagg cggataaagt 9060
tgcaggacca cttctgcgct cggcccttcc ggctggctgg tttattgctg ataaatctgg 9120
agccggtgag cgtgggtctc gcggtatcat tgcagcactg gggccagatg gtaagccctc 9180
ccgtatcgta gttatctaca cgacggggag tcaggcaact atggatgaac gaaatagaca 9240
gatcgctgag ataggtgcct cactgattaa gcattggtaa ctgtcagacc aagtttactc 9300
atatatactt tagattgatt taaaacttca tttttaattt aaaaggatct aggtgaagat 9360
cctttttgat aatctcatga ccaaaatccc ttaacgtgag ttttcgttcc actgagcgtc 9420
agaccccgta gaaaagatca aaggatcttc ttgagatcct ttttttctgc gcgtaatctg 9480
ctgcttgcaa acaaaaaaac caccgctacc agcggtggtt tgtttgccgg atcaagagct 9540
accaactctt tttccgaagg taactggctt cagcagagcg cagataccaa atactgtcct 9600
tctagtgtag ccgtagttag gccaccactt caagaactct gtagcaccgc ctacatacct 9660
cgctctgcta atcctgttac cagtggctgc tgccagtggc gataagtcgt gtcttaccgg 9720
gttggactca agacgatagt taccggataa ggcgcagcgg tcgggctgaa cggggggttc 9780
gtgcacacag cccagcttgg agcgaacgac ctacaccgaa ctgagatacc tacagcgtga 9840
gctatgagaa agcgccacgc ttcccgaagg gagaaaggcg gacaggtatc cggtaagcgg 9900
cagggtcgga acaggagagc gcacgaggga gcttccaggg ggaaacgcct ggtatcttta 9960
tagtcctgtc gggtttcgcc acctctgact tgagcgtcga tttttgtgat gctcgtcagg 10020
ggggcggagc ctatggaaaa acgccagcaa cgcggccttt ttacggttcc tggccttttg 10080
ctggcctttt gctcacatgg ctcgacagat ct 10112

<210> 54
<211> 10114
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic

oligonucleotide

<400> 54
tcaatattgg ccattagcca tattattcat tggttatata gcataaatca atattggcta 60
ttggccattg catacgttgt atctatatca taatatgtac atttatattg gctcatgtcc 120
aatatgaccg ccatgttggc attgattatt gactagttat taatagtaat caattacggg 180
gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc 240
gcctggctga ccgcccaacg acccccgccc attgacgtca ataatgacgt atgttcccat 300
agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc 360
ccacttggca gtacatcaag tgtatcatat gccaagtccg ccccctattg acgtcaatga 420
cggtaaatgg cccgcctggc attatgccca gtacatgacc ttacgggact ttcctacttg 480
gcagtacatc tacgtattag tcatcgctat taccatggtg atgcggtttt ggcagtacac 540
caatgggcgt ggatagcggt ttgactcacg gggatttcca agtctccacc ccattgacgt 600
caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactg 660
cgatcgcccg ccccgttgac gcaaatgggc ggtaggcgtg tacggtggga ggtctatata 720
agcagagctc gtttagtgaa ccgtcagatc actagaagct ttattgcggt agtttatcac 780
agttaaattg ctaacgcagt cagtgcttct gacacaacag tctcgaactt aagctgcagt 840
gactctctta aggtagcctt gcagaagttg gtcgtgaggc actgggcagg taagtatcaa 900
ggttacaaga caggtttaag gagaccaata gaaactgggc ttgtcgagac agagaagact 960
cttgcgtttc tgataggcac ctattggtct tactgacatc cactttgcct ttctctccac 1020
aggtgtccac tcccagttca attacagctc ttaaggctag agtacttaat acgactcact 1080
ataggctaga gaattccagg taagatgggc gatcccctca cctggtccaa agccctgaag 1140
aaactggaaa aagtcaccgt tcagggtagc caaaagctta ccacaggcaa ttgcaactgg 1200
gcattgtccc tggtggatct ttccacgac actaatttcg ttaaggagaa agattggcaa 1260
ctcagagacg tgatccccct cttggaggac gtgacccaaa cattgtctgg gcaggagcgc 1320
gaagctttcg agcgcacctg gtgggccatc agcgcagtca aaatggggct gcaaatcaac 1380
aacgtggttg acggtaaagc tagctttcaa ctgctccgcg ctaagtacga gaagaaaacc 1440
gccaacaaga aacaatccga acctagcgag gagtacccaa ttatgatcga cggcgccggc 1500
aataggaact tccgcccact gactcccagg ggctatacca cctgggtcaa caccatccag 1560
acaaacggac ttttgaacga agcctcccag aacctgttcg gcatcctgtc tgtggactgc 1620
acctccgaag aaatgaatgc ttttctcgac gtggtgccag gacaggctgg acagaaacag 1680
atcctgctcg atgccattga caagatcgcc gacgactggg ataatcgcca cccccctgcca 1740
aacgcccctc tggtggctcc cccacagggg cctatcccta tgaccgctag gttcattagg 1800
ggactggggg tgccccgcga acgccagatg gagccagcat ttgaccaatt taggcagacc 1860
tacagacagt ggatcatcga agccatgagc gaggggatta aagtcatgat cggaaagccc 1920
aaggcacaga acatcaggca gggggccaag gaaccatacc ctgagtttgt cgacaggctt 1980
ctgtcccaga ttaaatccga aggccacccct caggagatct ccaagttctt gacagacaca 2040
ctgactatcc aaaatgcaaa tgaagagtgc agaaacgcca tgaggcacct cagacctgaa 2100
gataccctgg aggagaaaat gtacgcatgt cgcgacattg gcactaccaa gcaaaagatg 2160
atgctgctcg ccaaggctct gcaaaccggc ctggctggtc cattcaaagg aggagcactg 2220
aagggaggtc cattgaaagc tgcacaaaca tgttataatt gtgggaagcc aggacattta 2280
tctagtcaat gtagagcacc taaagtctgt tttaaatgta aacagcctgg acatttctca 2340
aagcaatgca gaagtgttcc aaaaaacggg aagcaagggg ctcaagggag gccccagaaa 2400
caaactttcc cgatacaaca gaagagtcag cacaacaaat ctgttgtaca agagactcct 2460
cagactcaaa atctgtaccc agatctgagc gaaataaaaa aggaatacaa tgtcaaggag 2520
aaggatcaag tagaggatct caacctggac agtttgtggg agtaacatac aatctcgaga 2580
agaggcccac taccatcgtc ctgatcaatg acacccctct taatgtgctg ctggacaccg 2640
gagccgacac cagcgttctc actactgctc actataacag actgaaatac agaggaagga 2700
aataccaggg cacaggcatc atcggcgttg gaggcaacgt cgaaaccttt tccactcctg 2760
tcaccatcaa aaagaagggg agacacatta aaaccagaat gctggtcgcc gacatccccg 2820
tcaccatcct tggcagagac attctccagg acctgggcgc taaactcgtg ctggcacaac 2880
tgtctaagga aatcaagttc cgcaagatcg agctgaaaga gggcacaatg ggtccaaaaa 2940
tcccccagtg gcccctgacc aaagagaagc ttgagggcgc taaggaaatc gtgcagcgcc 3000
tgctttctga gggcaagatt agcgaggcca gcgacaataa cccttacaac agccccatct 3060
ttgtgattaa gaaaaggagc ggcaaatgga gactcctgca ggacctgagg gaactcaaca 3120
agaccgtcca ggtcggaact gagatctctc gcggactgcc tcaccccggc ggcctgatta 3180
aatgcaagca catgacagtc cttgacattg gagacgctta ttttaccatc cccctcgatc 3240
ctgaatttcg cccctatact gcttttacca tccccagcat caatcaccag gagcccgata 3300

aacgctatgt gtggaagtgc ctcccccagg gatttgtgct tagcccctac atttaccaga 3360
agacacttca agagatcctc caacctttcc gcgaaagata cccagaggtt caactctacc 3420
aatatatgga cgacctgttc atggggtcca acgggtctaa gaagcagcac aaggaactca 3480
tcatcgaact gagggcaatc ctcctggaga aaggcttcga gacacccgac gacaagctgc 3540
aagaagttcc tccatatagc tggctgggct accagctttg ccctgaaaac tggaaagtcc 3600
agaagatgca gttggatatg gtcaagaacc caacactgaa cgacgtccag aagctcatgg 3660
gcaatattac ctggatgagc tccggaatcc ctgggcttac cgttaagcac attgccgcaa 3720
ctacaaaagg atgcctggag ttgaaccaga aggtcatttg gacagaggaa gctcagaagg 3780
aactggagga gaataatgaa aagattaaga atgctcaagg gctccaatac tacaatcccg 3840
aagaagaaat gttgtgcgag gtcgaaatca ctaagaacta cgaagccacc tatgtcatca 3900
aacagtccca aggcatcttg tgggccggaa agaaaatcat gaaggccaac aaaggctggt 3960
ccaccgttaa aaatctgatg ctcctgctcc agcacgtcgc caccgagtct atcacccgcg 4020
tcggcaagtg ccccaccttc aaagttccct tcactaagga gcaggtgatg tgggagatgc 4080
aaaaaggctg gtactactct tggcttcccg agatcgtcta cacccaccaa gtggtgcacg 4140
acgactggag aatgaagctt gtcgaggagc ccactagcgg aattacaatc tataccgacg 4200
gcggaaagca aaacggagag ggaatcgctg catacgtcac atctaacggc cgcaccaagc 4260
aaaagaggct cggccctgtc actcaccagg tggctgagag gatggctatc cagatggccc 4320
ttgaggacac tagagacaag caggtgaaca ttgtgactga cagctactac tgctggaaaa 4380
acatcacaga gggccttggc ctggagggac cccagtctcc ctggtggcct atcatccaga 4440
atatccgcga aaaggaaatt gtctatttcg cctgggtgcc tggacacaaa ggaatttacg 4500
gcaaccaact cgccgatgaa gccgccaaaa ttaaagagga aatcatgctt gcctaccagg 4560
gcacacagat taaggagaag agagacgagg acgctggctt tgacctgtgt gtgccatacg 4620
acatcatgat tcccgttagc gacacaaaga tcattccaac cgatgtcaag atccaggtgc 4680
cacccaattc atttggttgg gtgaccggaa agtccagcat ggctaagcag ggtcttctga 4740
ttaacggggg aatcattgat gaaggataca ccggcgaaat ccaggtgatc tgcacaaata 4800
tcggcaaaag caatattaag cttatcgaag ggcagaagtt cgctcaactc atcatcctcc 4860
agcaccacag caattcaaga caaccttggg acgaaaacaa gattagccag agaggtgaca 4920
agggcttcgg cagcacaggt gtgttctggg tggagaacat ccaggaagca caggacgagc 4980
acgagaattg gcacacctcc cctaagattt tggcccgcaa ttacaagatc ccactgactg 5040
tggctaagca gatcacacag gaatgccccc actgcaccaa acaaggttct ggccccgccg 5100
gctgcgtgat gaggtccccc aatcactggc aggcagattg cacccacctc gacaacaaaa 5160
ttatcctgac cttcgtggag agcaattccg gctacatcca cgcaacactc ctctccaagg 5220
aaaatgcatt gtgcacctcc ctcgcaattc tggaatgggc caggctgttc tctccaaaat 5280
ccctgcacac cgacaacggc accaactttg tggctgaacc tgtggtgaat ctgctgaagt 5340
tcctgaaaat cgcccacacc actggcattc cctatcaccc tgaaagccag ggcattgtcg 5400
agagggccaa cagaactctg aaagaaaaga tccaatctca cagagacaat acacagacat 5460
tggaggccgc acttcagctc gcccttatca cctgcaacaa aggaagagaa agcatgggcg 5520
gccagacccc ctgggaggtc ttcatcacta accaggccca ggtcatccat gaaaagctgc 5580
tcttgcagca ggcccagtcc tccaaaaagt tctgctttta taagatcccc ggtgagcacg 5640
actggaaagg tcctacaaga gttttgtgga aaggagacgg cgcagttgtg gtgaacgatg 5700
agggcaaggg gatcatcgct gtgcccctga cacgcaccaa gcttctcatc aagccaaact 5760
gaacccgggg cggccgcttc cctttagtga gggttaatgc ttcgagcaga catgataaga 5820
tacattgatg agtttggaca aaccacaact agaatgcagt gaaaaaaatg ctttatttgt 5880
gaaatttgtg atgctattgc tttatttgta accattataa gctgcaataa acaagttaac 5940
aacaacaatt gcattcattt tatgtttcag gttcaggggg agatgtggga ggttttttaa 6000
agcaagtaaa acctctacaa atgtggtaaa atccgataag gatcgatccg ggctggcgta 6060
atagcgaaga ggcccgcacc gatcgccctt cccaacagtt gcgcagcctg aatggcgaat 6120
ggacgcgccc tgtagcggcg cattaagcgc ggcgggtgtg gtggttacgc gcagcgtgac 6180
cgctacactt gccagcgccc tagcgcccgc tcctttcgct ttcttccctt cctttctcgc 6240
cacgttcgcc ggctttcccc gtcaagctct aaatcggggg ctccctttag ggttccgatt 6300
tagagcttta cggcacctcg accgcaaaaa acttgatttg ggtgatggtt cacgtagtgg 6360
gccatcgccc tgatagacgg tttttcgccc tttgacgttg gagtccacgt tctttaatag 6420
tggactcttg ttccaaactg gaacaacact caaccctatc tcggtctatt cttttgattt 6480
ataagggatt ttgccgattt cggcctattg gttaaaaaat gagctgattt aacaaatatt 6540
taacgcgaat tttaacaaaa tattaacgtt tacaatttcg cctgatgcgg tattttctcc 6600
ttacgcatct gtgcggtatt tcacaccgca tacgcggatc tgcgcagcac catggcctga 6660
aataacctct gaaagaggaa cttggttagg taccttctga ggcggaaaga accagctgtg 6720
gaatgtgtgt cagttagggt gtggaaagtc cccaggctcc ccagcaggca gaagtatgca 6780

aagcatgcat ctcaattagt cagcaaccag gtgtggaaag tccccaggct ccccagcagg 6840
cagaagtatg caaagcatgc atctcaatta gtcagcaacc atagtcccgc ccctaactcc 6900
gcccatcccg cccctaactc cgcccagttc cgcccattct ccgccccatg gctgactaat 6960
ttttttattt tatgcagagg ccgaggccgc ctcggcctct gagctattcc agaagtagtg 7020
aggaggcttt tttggaggcc taggcttttg caaaaagctt gattcttctg acacaacagt 7080
ctcgaactta aggctagagc caccatgatt gaacaagatg gattgcacgc aggttctccg 7140
gccgcttggg tggagaggct attcggctat gactgggcac aacagacaat cggctgctct 7200
gatgccgccg tgttccggct gtcagcgcag gggcgcccgg ttctttttgt caagaccgac 7260
ctgtccggtg ccctgaatga actgcaggac gaggcagcgc ggctatcgtg gctggccacg 7320
acgggcgttc cttgcgcagc tgtgctcgac gttgtcactg aagcgggaag ggactggctg 7380
ctattgggcg aagtgccggg gcaggatctc ctgtcatctc accttgctcc tgccgagaaa 7440
gtatccatca tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc tacctgccca 7500
ttcgaccacc aagcgaaaca tcgcatcgag cgagcacgta ctcggatgga agccggtctt 7560
gtcgatcagg atgatctgga cgaagagcat caggggctcg cgccagccga actgttcgcc 7620
aggctcaagg cgcgcatgcc cgacggcgag gatctcgtcg tgacccatgg cgatgcctgc 7680
ttgccgaata tcatggtgga aaatggccgc ttttctggat tcatcgactg tggccggctg 7740
ggtgtggcgg accgctatca ggacatagcg ttggctaccc gtgatattgc tgaagagctt 7800
ggcggcgaat gggctgaccg cttcctcgtg ctttacggta tcgccgctcc cgattcgcag 7860
cgcatcgcct tctatcgcct tcttgacgag ttcttctgag cgggactctg gggttcgaaa 7920
tgaccgacca agcgacgccc aacctgccat cacgatggcc gcaataaaat atctttattt 7980
tcattacatc tgtgtgttgg ttttttgtgt gaatcgatag cgataaggat ccgcgtatgg 8040
tgcactctca gtacaatctg ctctgatgcc gcatagttaa gccagccccg acacccgcca 8100
acacccgctg acgcgccctg acgggcttgt ctgctcccgg catccgctta cagacaagct 8160
gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg 8220
agacgaaagg gcctcgtgat acgcctattt ttataggtta atgtcatgat aataatggtt 8280
tcttagacgt caggtggcac ttttcgggga aatgtgcgcg gaacccctat ttgtttattt 8340
ttctaaatac attcaaatat gtatccgctc atgagacaat aaccctgata aatgcttcaa 8400
taatattgaa aaaggaagag tatgagtatt caacatttcc gtgtcgccct tattcccttt 8460
tttgcggcat tttgccttcc tgtttttgct cacccagaaa cgctggtgaa agtaaaagat 8520
gctgaagatc agttgggtgc acgagtgggt tacatcgaac tggatctcaa cagcggtaag 8580
atccttgaga gttttcgccc cgaagaacgt tttccaatga tgagcacttt taaagttctg 8640
ctatgtggcg cggtattatc ccgtattgac gccgggcaag agcaactcgg tcgccgcata 8700
cactattctc agaatgactt ggttgagtac tcaccagtca cagaaaagca tcttacggat 8760
ggcatgacag taagagaatt atgcagtgct gccataacca tgagtgataa cactgcggcc 8820
aacttacttc tgacaacgat cggaggaccg aaggagctaa ccgctttttt gcacaacatg 8880
ggggatcatg taactcgcct tgatcgttgg gaaccggagc tgaatgaagc cataccaaac 8940
gacgagcgtg acaccacgat gcctgtagca atggcaacaa cgttgcgcaa actattaact 9000
ggcgaactac ttactctagc ttcccggcaa caattaatag actggatgga ggcggataaa 9060
gttgcaggac cacttctgcg ctcggccctt ccggctggct ggtttattgc tgataaatct 9120
ggagccggtg agcgtgggtc tcgcggtatc attgcagcac tggggccaga tggtaagccc 9180
tcccgtatcg tagttatcta cacgacgggg agtcaggcaa ctatggatga acgaaataga 9240
cagatcgctg agataggtgc ctcactgatt aagcattggt aactgtcaga ccaagtttac 9300
tcatatatac tttagattga tttaaaactt catttttaat ttaaaaggat ctaggtgaag 9360
atcctttttg ataatctcat gaccaaaatc ccttaacgtg agttttcgtt ccactgagcg 9420
tcagaccccg tagaaaagat caaaggatct tcttgagatc ctttttttct gcgcgtaatc 9480
tgctgcttgc aaacaaaaaa accaccgcta ccagcggtgg tttgtttgcc ggatcaagag 9540
ctaccaactc tttttccgaa ggtaactggc ttcagcagag cgcagatacc aaatactgtc 9600
cttctagtgt agccgtagtt aggccaccac ttcaagaact ctgtagcacc gcctacatac 9660
ctcgctctgc taatcctgtt accagtggct gctgccagtg gcgataagtc gtgtcttacc 9720
gggttggact caagacgata gttaccggat aaggcgcagc ggtcgggctg aacggggggt 9780
tcgtgcacac agcccagctt ggagcgaacg acctacaccg aactgagata cctacagcgt 9840
gagctatgag aaagcgccac gcttcccgaa gggagaaagg cggacaggta tccggtaagc 9900
ggcagggtcg gaacaggaga gcgcacgagg gagcttccag ggggaaacgc ctggtatctt 9960
tatagtcctg tcgggtttcg ccacctctga cttgagcgtc gatttttgtg atgctcgtca 10020
gggggcgga gcctatggaa aaacgccagc aacgcggcct ttttacggtt cctggccttt 10080
tgctggcctt ttgctcacat ggctcgacag atct 10114

<210> 55
<211> 12473
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 55
gttaccttct gctctgcaga atggccaacc tttaacgtcg gatggccgcg agacggcacc 60
tttaaccgag acctcatcac ccaggttaag atcaaggtct tttcacctgg cccgcatgga 120
cacccagacc aggtccccta catcgtgacc tgggaagcct tggcttttga cccccctccc 180
tgggtcaagc cctttgtaca ccctaagcct ccgcctcctc ttcctccatc cgccccgtct 240
ctccccccttg aacctcctcg ttcgaccccg cctcgatcct ccctttatcc agccctcact 300
ccttctctag gcgccggaat tcgttaactc gagaggcctg ccaccatggg gactgctcca 360
aagaagaagc gtaaggtagt cgttttacaa cgtcgtgact gggaaaaccc tggcgttacc 420
caacttaatc gccttgcagc acatccccct ttcgccagct ggcgtaatag cgaagaggcc 480
cgcaccgatc gcccttccca acagttgcgc agcctgaatg gcgaatggcg ctttgcctgg 540
tttccggcac cagaagcggt gccggaaagc tggctggagt gcgatcttcc tgaggccgat 600
actgtcgtcg tcccctcaaa ctggcagatg cacggttacg atgcgcccat ctacaccaac 660
gtaacctatc ccattacggt caatccgccg tttgttccca cggagaatcc gacgggttgt 720
tactcgctca catttaatgt tgatgaaagc tggctacagg aaggccagac gcgaattatt 780
tttgatggcg ttaactcggc gtttcatctg tggtgcaacg ggcgctgggt cggttacggc 840
caggacagtc gtttgccgtc tgaatttgac ctgagcgcat ttttacgcgc cggagaaaac 900
cgcctcgcgg tgatggtgct gcgttggagt gacggcagtt atctggaaga tcaggatatg 960
tggcggatga gcggcatttt ccgtgacgtc tcgttgctgc ataaaccgac tacacaaatc 1020
agcgatttcc atgttgccac tcgctttaat gatgatttca gccgcgctgt actggaggct 1080
gaagttcaga tgtgcggcga gttgcgtgac tacctacggg taacagtttc tttatggcag 1140
ggtgaaacgc aggtcgccag cggcaccgcg cctttcggcg gtgaaattat cgatgagcgt 1200
ggtggttatg ccgatcgcgt cacactacgt ctgaacgtcg aaaacccgaa actgtggagc 1260
gccgaaatcc cgaatctcta tcgtgcggtg gttgaactgc acaccgccga cggcacgctg 1320
attgaagcag aagcctgcga tgtcggtttc cgcgaggtgc ggattgaaaa tggtctgctg 1380
ctgctgaacg gcaagccgtt gctgattcga ggcgttaacc gtcacgagca tcatcctctg 1440
catggtcagg tcatggatga gcagacgatg gtgcaggata tcctgctgat gaagcagaac 1500
aactttaacg ccgtgcgctg ttcgcattat ccgaaccatc cgctgtggta cacgctgtgc 1560
gaccgctacg gcctgtatgt ggtggatgaa gccaatattg aaacccacgg catggtgcca 1620
atgaatcgtc tgaccgatga tccgcgctgg ctaccggcga tgagcgaacg cgtaacgcga 1680
atggtgcagc gcgatcgtaa tcacccgagt gtgatcatct ggtcgctggg gaatgaatca 1740
ggccacggcg ctaatcacga cgcgctgtat cgctggatca aatctgtcga tccttcccgc 1800
ccggtgcagt atgaaggcgg cggagccgac accacggcca ccgatattat ttgcccgatg 1860
tacgcgcgcg tggatgaaga ccagcccttc ccggctgtgc cgaaatggtc catcaaaaaa 1920
tggctttcgc tacctggaga gacgcgcccg ctgatccttt gcgaatacgc ccacgcgatg 1980
ggtaacagtc ttggcggttt cgctaaatac tggcaggcgt ttcgtcagta tccccgttta 2040
cagggcggct tcgtctggga ctgggtggat cagtcgctga ttaaatatga tgaaaacggc 2100
aacccgtggt cggcttacgg cggtgatttt ggcgatacgc cgaacgatcg ccagttctgt 2160
atgaacggtc tggtctttgc cgaccgcacg ccgcatccag cgctgacgga agcaaaacac 2220
cagcagcagt ttttccagtt ccgtttatcc gggcaaacca tcgaagtgac cagcgaatac 2280
ctgttccgtc atagcgataa cgagctcctg cactggatgg tggcgctgga tggtaagccg 2340
ctggcaagcg gtgaagtgcc tctggatgtc gctccacaag gtaaacagtt gattgaactg 2400
cctgaactac cgcagccgga gagcgccggg caactctggc tcacagtacg cgtagtgcaa 2460
ccgaacgcga ccgcatggtc agaagccggg cacatcagcg cctggcagca gtggcgtctg 2520
gcggaaaacc tcagtgtgac gctccccgcc gcgtcccacg ccatcccgca tctgaccacc 2580
agcgaaatgg atttttgcat cgagctgggt aataagcgtt ggcaatttaa ccgccagtca 2640
ggctttcttt cacagatgtg gattggcgat aaaaaacaac tgctgacgcc gctgcgcgat 2700
cagttcaccc gtgcaccgct ggataacgac attggcgtaa gtgaagcgac ccgcattgac 2760
cctaacgcct gggtcgaacg ctggaaggcg gcgggccatt accaggccga agcagcgttg 2820
ttgcagtgca cggcagatac acttgctgat gcggtgctga ttacgaccgc tcacgcgtgg 2880

cagcatcagg ggaaaacctt atttatcagc cggaaaacct accggattga tggtagtggt 2940
caaatggcga ttaccgttga tgttgaagtg gcgagcgata caccgcatcc ggcgcggatt 3000
ggcctgaact gccagctggc gcaggtagca gagcgggtaa actggctcgg attagggccg 3060
caagaaaact atcccgaccg ccttactgcc gcctgttttg accgctggga tctgccattg 3120
tcagacatgt ataccccgta cgtcttcccg agcgaaaacg gtctgcgctg cgggacgcgc 3180
gaattgaatt atggcccaca ccagtggcgc ggcgacttcc agttcaacat cagccgctac 3240
agtcaacagc aactgatgga aaccagccat cgccatctgc tgcacgcgga agaaggcaca 3300
tggctgaata tcgacggttt ccatatgggg attggtggcg acgactcctg gagcccgtca 3360
gtatcggcgg aattccagct gagcgccggt cgctaccatt accagttggt ctggtgtcaa 3420
aaataataat aaccgggcag gggggatccg cagatccggc tgtggaatgt gtgtcagtta 3480
gggtgtggaa agtccccagg ctccccagca ggcagaagta tgcaaagcat gcctgcagga 3540
gtggggaggc acgatggccg ctttggtcga ggcggatccg gccattagcc atattattca 3600
ttggttatat agcataaatc aatattggct attggccatt gcatacgttg tatccatatc 3660
ataatatgta catttatatt ggctcatgtc caacattacc gccatgttga cattgattat 3720
tgactagtta ttaatagtaa tcaattacgg ggtcattagt tcatagccca tatatggagt 3780
tccgcgttac ataacttacg gtaaatggcc cgcctggctg accgcccaac gacccccgcc 3840
cattgacgtc aataatgacg tatgttccca tagtaacgcc aatagggact ttccattgac 3900
gtcaatgggt ggagtattta cggtaaactg cccacttggc agtacatcaa gtgtatcata 3960
tgccaagtac gccccctatt gacgtcaatg acggtaaatg gcccgcctgg cattatgccc 4020
agtacatgac cttatgggac tttcctactt ggcagtacat ctacgtatta gtcatcgcta 4080
ttaccatggt gatgcggttt tggcagtaca tcaatgggcg tggatagcgg tttgactcac 4140
ggggatttcc aagtctccac cccattgacg tcaatgggag tttgttttgg caccaaaatc 4200
aacgggactt tccaaaatgt cgtaacaact ccgccccatt gacgcaaatg ggcggtaggc 4260
atgtacggtg ggaggtctat ataagcagag ctcgtttagt gaaccgtcag atcgcctgga 4320
gacgccatcc acgctgtttt gacctccata gaagacaccg ggaccgatcc agcctccgcg 4380
gccccaagct tgttgggatc caccggtcgc caccatggtg agcaagggcg aggagctgtt 4440
caccggggtg gtgcccatcc tggtcgagct ggacggcgac gtaaacggcc acaagttcag 4500
cgtgtccggc gagggcgagg gcgatgccac ctacggcaag ctgaccctga agttcatctg 4560
caccaccggc aagctgcccg tgccctggcc caccctcgtg accaccctga cctacggcgt 4620
gcagtgcttc agccgctacc ccgaccacat gaagcagcac gacttcttca agtccgccat 4680
gcccgaaggc tacgtccagg agcgcaccat cttcttcaag gacgacggca actacaagac 4740
ccgcgccgag gtgaagttcg agggcgacac cctggtgaac cgcatcgagc tgaagggcat 4800
cgacttcaag gaggacggca acatcctggg gcacaagctg gagtacaact acaacagcca 4860
caacgtctat atcatggccg acaagcagaa gaacggcatc aaggtgaact tcaagatccg 4920
ccacaacatc gaggacggca gcgtgcagct cgccgaccac taccagcaga acacccccat 4980
cggcgacggc cccgtgctgc tgcccgacaa ccactacctg agcacccagt ccgccctgag 5040
caaagacccc aacgagaagc gcgatcacat ggtcctgctg gagttcgtga ccgccgccgg 5100
gatcactctc ggcatggacg agctgtacaa gtaaagcggc cgcgactcta gatcataatc 5160
agccatacca catttgtaga ggttttactt gctttaaaaa acctcccaca cctccccctg 5220
aacctgaaac ataaaatgaa tgcaattgtt gttgttaaca tcgataaaat aaaagatttt 5280
atttagtctc cagaaaaagg ggggaatgaa agaccccacc tgtaggtttg gcaagctagc 5340
ttaagtaacg ccattttgca aggcatggaa aaatacataa ctgagaatag agaagttcag 5400
atcaaggtca ggaacagatg gaacagctga atatgggcca aacaggatat ctgtggtaag 5460
cagttcctgc cccggctcag ggccaagaac agatggaaca gctgaatatg ggccaaacag 5520
gatatctgtg gtaagcagtt cctgccccgg ctcagggcca agaacagatg gtccccagat 5580
gcggtccagc cctcagcagt ttctagagaa ccatcagatg tttccagggt gccccaagga 5640
cctgaaatga ccctgtgcct tatttgaact aaccaatcag ttcgcttctc gcttctgttc 5700
gcgcgcttct gctccccgag ctcaataaaa gagcccacaa cccctcactc ggggcgccag 5760
tcctccgatt gactgagtcg cccgggtacc cgtgtatcca ataaaccctc ttgcagttgc 5820
atccgacttg tggtctcgct gttccttggg agggtctcct ctgagtgatt gactacccgt 5880
cagcgggggt ctttcatttg ggggctcgtc cgggatcggg agacccctgc ccagggacca 5940
ccgacccacc accgggaggt aagctggcca gcaacttatc tgtgtctgtc cgattgtcta 6000
gtgtctatga ctgattttat gcgcctgcgt cggtactagt tagctaacta gctctgtatc 6060
tggcggaccc gtggtggaac tgacgagttc ggaacacccg gccgcaaccc tgggagagga 6120
attctcatgt ttgacagctt atcatcgata agctttttgc aaaagcctag gcctccaaaa 6180
aagcctcctc actacttctg gaatagctca gaggccgagg cggcctcggc ctctgcataa 6240
ataaaaaaaa ttagtcagcc atggggcgga gaatgggcgg aactgggcgg agttaggggc 6300
gggatgggcg gagttagggg cgggactatg gttgctgact aattgagatg catgctttgc 6360

atacttctgc ctgctgggga gcctggggac tttccacacc tggttgctga ctaattgaga 6420
tgcatgcttt gcatacttct gcctgctggg gagcctgggg actttccaca ccctaactga, 6480
cacacattcc acagccggat cctctacgcc ggacgcatcg tggccggcat caccggcgcc 6540
acaggtgcgg ttgctggcgc ctatatcgcc gacatcaccg atggggaaga tcgggctcgc 6600
cacttcgggc tcatgagcgc ttgtttcggc gtgggtatgg tggcaggccc cgtggccggg 6660
ggactgttgg gcgccatctc cttgcatgca ccattccttg cggcggcggt gctcaacggc 6720
ctcaacctac tactgggctg cttcctaatg caggagtcgc ataagggaga gcgtcgaccg 6780
atgcccttga gagccttcaa cccagtcagc tccttccggt gggcgcgggg catgactatc 6840
gtcgccgcac ttatgactgt cttctttatc atgcaactcg taggacaggt gccggcagcg 6900
ctctgggtca ttttcggcga ggaccgcttt cgctggagcg cgacgatgat cggcctgtcg 6960
cttgcggtat tcggaatctt gcacgccctc gctcaagcct tcgtcactgg tcccgccacc 7020
aaacgtttcg gcgagaagca ggccattatc gccggcatgg cggccgacgc gctgggctac 7080
gtcttgctgg cgttcgcgac gcgaggctgg atggccttcc ccattatgat tcttctcgct 7140
tccggcggca tcgggatgcc cgcgttgcag gccatgctgt ccaggcaggt agatgacgac 7200
catcagggac agcttcaagg atcgctcgcg gctcttacca gcctaacttc gatcactgga 7260
ccgctgatcg tcacggcgat ttatgccgcc tcggcgagca catggaacgg gttggcatgg 7320
attgtaggcg ccgccctata ccttgtctgc ctccccgcgt tgcgtcgcgg tgcatggagc 7380
cgggccacct cgacctgaat ggaagccggc ggcacctcgc taacggattc accactccaa 7440
gaattggagc caatcaattc ttgcggagaa ctgtgaatgc gcaaaccaac ccttggcaga 7500
acatatccat cgcgtccgcc atctccagca gccgcacgcg gcgcatctcg ggcagcgttg 7560
ggtcctggcc acgggtgcgc atgatcgtgc tcctgtcgtt gaggacccgg ctaggctggc 7620
ggggttgcct tactggttag cagaatgaat caccgatacg cgagcgaacg tgaagcgact 7680
gctgctgcaa aacgtctgcg acctgagcaa caacatgaat ggtcttcggt ttccgtgttt 7740
cgtaaagtct ggaaacgcgg aagtcagcgc cctgcaccat tatgttccgg atctgcatcg 7800
caggatgctg ctggctaccc tgtggaacac ctacatctgt attaacgaag cgctggcatt 7860
gaccctgagt gatttttctc tggtcccgcc gcatccatac cgccagttgt ttaccctcac 7920
aacgttccag taaccgggca tgttcatcat cagtaacccg tatcgtgagc atcctctctc 7980
gtttcatcgg tatcattacc cccatgaaca gaaattcccc cttacacgga ggcatcaagt 8040
gaccaaacag gaaaaaaccg cccttaacat ggcccgcttt atcagaagcc agacattaac 8100
gcttctggag aaactcaacg agctggacgc ggatgaacag gcagacatct gtgaatcgct 8160
tcacgaccac gctgatgagc tttaccgcag ctgcctcgcg cgtttcggtg atgacggtga 8220
aaacctctga cacatgcagc tcccggagac ggtcacagct tgtctgtaag cggatgccgg 8280
gagcagacaa gcccgtcagg gcgcgtcagc gggtgttggc gggtgtcggg gcgcagccat 8340
gacccagtca cgtagcgata gcggagtgta tactggctta actatgcggc atcagagcag 8400
attgtactga gagtgcacca tatgcggtgt gaaataccgc acagatgcgt aaggagaaaa 8460
taccgcatca ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc ggtcgttcgg 8520
ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac agaatcaggg 8580
gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa ccgtaaaaag 8640
gccgcgttgc tggcgttttt ccataggctc cgcccccctg acgagcatca caaaaatcga 8700
cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc gtttccccct 8760
ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata cctgtccgcc 8820
tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta tctcagttcg 8880
gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca gcccgaccgc 8940
tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga cttatcgcca 9000
ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg tgctacagag 9060
ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg tatctgcgct 9120
ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg caaacaaacc 9180
accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag aaaaaaagga 9240
tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa cgaaaactca 9300
cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat ccttttaaat 9360
taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac 9420
caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt 9480
gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt 9540
gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc aataaaccag 9600
ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc catccagtct 9660
attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacgtt 9720
gttgccattg ctgcaggcat cgtggtgtca cgctcgtcgt ttggtatggc ttcattcagc 9780
tccggttccc aacgatcaag gcgagttaca tgatccccca tgttgtgcaa aaaagcggtt 9840

agctccttcg gtcctccgat cgtcgttgtc agaagtaagt tggccgcagt gttatcactc 9900
atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct 9960
gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc 10020
tcttgcccgg cgtcaatacg ggataatacc gcgccacata gcagaacttt aaaagtgctc 10080
atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc 10140
agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc 10200
gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca 10260
cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt 10320
tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt 10380
ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca 10440
ttaacctata aaaataggcg tatcacgagg ccctttcgtc tcgcgcgttt cggtgatgac 10500
ggtgaaaacc tctgacacat gcagctcccg gagacggtca cagcttgtct gtaagcggat 10560
gccgggagca gacaagcccg tcagggcgcg tcagcgggtg ttggcgggtg tcggggctgg 10620
cttaactatg cggcatcaga gcagattgta ctgagagtgc accatatcga cgctctccct 10680
tatgcgactc ctgcattagg aagcagccca gtagtaggtt gaggccgttg agcaccgccg 10740
ccgcaaggaa tggtgcatgc aaggagatgg cgcccaacag tcccccggcc acggggcctg 10800
ccaccatacc cacgccgaaa caagcgctca tgagcccgaa gtggcgagcc cgatcttccc 10860
catcggtgat gtcggcgata taggcgccag caaccgcacc tgtggcgccg gtgatgccgg 10920
ccacgatgcg tccggcgtag aggatctggc tagcgatgac cctgctgatt ggttcgctga 10980
ccatttccgg ggtgcggaac ggcgttacca gaaactcaga aggttcgtcc aaccaaaccg 11040
actctgacgg cagtttacga gagagatgat agggtctgct tcagtaagcc agatgctaca 11100
caattaggct tgtacatatt gtcgttagaa cgcggctaca attaatacat aaccttatgt 11160
atcatacaca tacgatttag gtgacactat agaatacaag ctggaagatc ttccagcttg 11220
ggctgcaggt cgactctaga gtccgttaca taacttacgg taaatggccc gcctggctga 11280
ccgcccaacg acccccgccc attgacgtca ataatgacgt atgttcccat agtaacgcca 11340
atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc ccacttggca 11400
gtacatcaag tgtatcatat gccaagtacg ccccctattg acgtcaatga cggtaaatgg 11460
cccgcctggc attatgccca gtacatgacc ttatgggact ttcctacttg gcagtacatc 11520
tacgtattag tcatcgctat taccatggtg atgcggtttt ggcagtacat caatgggcgt 11580
ggatagcggt ttgactcacg gggatttcca agtctccacc ccattgacgt caatgggagt 11640
ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc cgccccattg 11700
acgcaaatgg gcggtaggcg tgtacggtgg gaggtctata taagcagagc tcgtttagtg 11760
aaccgcgcca gtcttccgat agactgcgtc gcccgggtac ccgtattccc aataaagcct 11820
cttgctgttt gcatccgaat cgtggtctcg ctgttccttg ggagggtctc ctctgagtga 11880
ttgactaccc acgacggggg tctttcattt gggggctcgt ccgggatttg gagacccctg 11940
cccagggacc accgacccac caccgggagg taagctggcc agcaacttat ctgtgtctgt 12000
ccgattgtct agtgtctatg tttgatgtta tgcgcctgcg tctgtactag ttagctaact 12060
agctctgtat ctggcggacc cgtggtggaa ctgacgagtt ctgaacaccc ggccgcaacc 12120
ctgggagacg tcccagggac tttgggggcc gtttttgtgg cccgacctga ggaagggagt 12180
cgatgtggaa tccgaccccg tcaggatatg tggttctggt aggagacgag aacctaaaac 12240
agttcccgcc tccgtctgaa tttttgcttt cggtttggaa ccgaagccgc gcgtcttgtc 12300
tgctgcagcg ctgcagcatc gttctgtgtt gtctctgtct gactgtgttt ctgtatttgt 12360
ctgaaaatta gggccagact gttaccactc ccttaagttt gaccttaggt cactggaaag 12420
atgtcgagcg gatcgctcac aaccagtcgg tagatgtcaa gaagagacgt tgg 12473

<210> 56
<211> 9729
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 56
gttaccttct gctctgcaga atggccaacc tttaacgtcg gatggccgcg agacggcacc 60
tttaaccgag acctcatcac ccaggttaag atcaaggtct tttcacctgg cccgcatgga 120

```
cacccagacc aggtccccta catcgtgacc tgggaagcct tggcttttga cccccctccc   180
tgggtcaagc cctttgtaca ccctaagcct ccgcctcctc ttcctccatc cgccccgtct   240
ctccccttg aacctcctcg ttcgaccccg cctcgatcct ccctttatcc agccctcact   300
ccttctctag gcgccggaat tcgttaactc gagaggcctg ccaccatggg gactgctcca   360
aagaagaagc gtaaggtagt cgttttacaa cgtcgtgact gggaaaaccc tggcgttacc   420
caacttaatc gccttgcagc acatccccct ttcgccagct ggcgtaatag cgaagaggcc   480
cgcaccgatc gcccttccca acagttgcgc agcctgaatg gcgaatggcg ctttgcctgg   540
tttccggcac cagaagcggt gccggaaagc tggctggagt gcgatcttcc tgaggccgat   600
actgtcgtcg tcccctcaaa ctggcagatg cacggttacg atgcgcccat ctacaccaac   660
gtaacctatc ccattacggt caatccgccg tttgttccca cggagaatcc gacgggttgt   720
tactcgctca catttaatgt tgatgaaagc tggctacagg aaggccagac gcgaattatt   780
tttgatggcg ttaactcggc gtttcatctg tggtgcaacg ggcgctgggt cggttacggc   840
caggacagtc gtttgccgtc tgaatttgac ctgagcgcat ttttacgcgc cggagaaaac   900
cgcctcgcgg tgatggtgct gcgttggagt gacggcagtt atctggaaga tcaggatatg   960
tggcggatga gcggcatttt ccgtgacgtc tcgttgctgc ataaaccgac tacacaaatc   1020
agcgatttcc atgttgccac tcgctttaat gatgatttca gccgcgctgt actggaggct   1080
gaagttcaga tgtgcggcga gttgcgtgac tacctacggg taacagtttc tttatggcag   1140
ggtgaaacgc aggtcgccag cggcaccgcg cctttcggcg gtgaaattat cgatgagcgt   1200
ggtggttatg ccgatcgcgt cacactacgt ctgaacgtcg aaaacccgaa actgtggagc   1260
gccgaaatcc cgaatctcta tcgtgcggtg gttgaactgc acaccgccga cggcacgctg   1320
attgaagcag aagcctgcga tgtcggtttc cgcgaggtgc ggattgaaaa tggtctgctg   1380
ctgctgaacg gcaagccgtt gctgattcga ggcgttaacc gtcacgagca tcatcctctg   1440
catggtcagg tcatggatga gcagacgatg gtgcaggata tcctgctgat gaagcagaac   1500
aactttaacg ccgtgcgctg ttcgcattat ccgaaccatc cgctgtggta cacgctgtgc   1560
gaccgctacg gcctgtatgt ggtggatgaa gccaatattg aaacccacgg catggtgcca   1620
atgaatcgtc tgaccgatga tccgcgctgg ctaccggcga tgagcgaacg cgtaacgcga   1680
atggtgcagc gcgatcgtaa tcacccgagt gtgatcatct ggtcgctggg gaatgaatca   1740
ggccacggcg ctaatcacga cgcgctgtat cgctggatca aatctgtcga tccttcccgc   1800
ccggtgcagt atgaaggcgg cggagccgac accacggcca ccgatattat ttgcccgatg   1860
tacgcgcgcg tggatgaaga ccagcccttc ccggctgtgc cgaaatggtc catcaaaaaa   1920
tggctttcgc tacctggaga gacgcgcccg ctgatccttt gcgaatacgc ccacgcgatg   1980
ggtaacagtc ttggcggttt cgctaaatac tggcaggcgt ttcgtcagta tccccgttta   2040
cagggcggct tcgtctggga ctgggtggat cagtcgctga ttaaatatga tgaaaacggc   2100
aacccgtggt cggcttacgg cggtgatttt ggcgatacgc cgaacgatcg ccagttctgt   2160
atgaacggtc tggtctttgc cgaccgcacg ccgcatccag cgctgacgga agcaaaacac   2220
cagcagcagt ttttccagtt ccgtttatcc gggcaaacca tcgaagtgac cagcgaatac   2280
ctgttccgtc atagcgataa cgagctcctg cactggatgg tggcgctgga tggtaagccg   2340
ctggcaagcg gtgaagtgcc tctggatgtc gctccacaag gtaaacagtt gattgaactg   2400
cctgaactac cgcagccgga gagcgccggg caactctggc tcacagtacg cgtagtgcaa   2460
ccgaacgcga ccgcatggtc agaagccggg cacatcagcg cctggcagca gtggcgtctg   2520
gcggaaaacc tcagtgtgac gctccccgcc gcgtcccacg ccatcccgca tctgaccacc   2580
agcgaaatgg atttttgcat cgagctgggt aataagcgtt ggcaatttaa ccgccagtca   2640
ggctttcttt cacagatgtg gattggcgat aaaaaacaac tgctgacgcc gctgcgcgat   2700
cagttcaccc gtgcaccgct ggataacgac attggcgtaa gtgaagcgac ccgcattgac   2760
cctaacgcct gggtcgaacg ctggaaggcg gcgggccatt accaggccga agcagcgttg   2820
ttgcagtgca cggcagatac acttgctgat gcggtgctga ttacgaccgc tcacgcgtgg   2880
cagcatcagg ggaaaacctt atttatcagc cggaaaacct accggattga tggtagtggt   2940
caaatggcga ttaccgttga tgttgaagtg gcgagcgata caccgcatcc ggcgcggatt   3000
ggcctgaact gccagctggc gcaggtagca gagcgggtaa actggctcgg attagggccg   3060
caagaaaact atcccgaccg ccttactgcc gcctgttttg accgctggga tctgccattg   3120
tcagacatgt ataccccgta cgtcttcccg agcgaaaacg gtctgcgctg cgggacgcgc   3180
gaattgaatt atggcccaca ccagtggcgc ggcgacttcc agttcaacat cagccgctac   3240
agtcaacagc aactgatgga aaccagccat cgccatctgc tgcacgcgga agaaggcaca   3300
tggctgaata tcgacggttt ccatatgggg attggtggcg acgactcctg gagcccgtca   3360
gtatcggcgg aattccagct gagcgccggt cgctaccatt accagttggt ctggtgtcaa   3420
aaataataat aaccgggcag gggggatccg cagatccggc tgtggaatgt gtgtcagtta   3480
gggtgtggaa agtccccagg ctccccagca ggcagaagta tgcaaagcat gcctgcagga   3540
gtggggaggc acgatggccg ctttggtcga ggcggatccg gccattagcc atattattca   3600
```

```
ttggttatat agcataaatc aatattggct attggccatt gcatacgttg tatccatatc 3660
ataatatgta catttatatt ggctcatgtc caacattacc gccatgttga cattgattat 3720
tgactagtta ttaatagtaa tcaattacgg ggtcattagt tcatagccca tatatggagt 3780
tccgcgttac ataacttacg gtaaatggcc cgcctggctg accgcccaac gacccccgcc 3840
cattgacgtc aataatgacg tatgttccca tagtaacgcc aatagggact ttccattgac 3900
gtcaatgggt ggagtattta cggtaaactg cccacttggc agtacatcaa gtgtatcata 3960
tgccaagtac gccccctatt gacgtcaatg acggtaaatg gcccgcctgg cattatgccc 4020
agtacatgac cttatgggac tttcctactt ggcagtacat ctacgtatta gtcatcgcta 4080
ttaccatggt gatgcggttt tggcagtaca tcaatgggcg tggatagcgg tttgactcac 4140
ggggatttcc aagtctccac cccattgacg tcaatgggag tttgttttgg caccaaaatc 4200
aacgggactt tccaaaatgt cgtaacaact ccgccccatt gacgcaaatg ggcggtaggc 4260
atgtacggtg ggaggtctat ataagcagag ctcgtttagt gaaccgtcag atcgcctgga 4320
gacgccatcc acgctgtttt gacctccata gaagacaccg ggaccgatcc agcctccgcg 4380
gccccaagct tgttgggatc caccggtcgc caccatggtg agcaagggcg aggagctgtt 4440
caccggggtg gtgcccatcc tggtcgagct ggacggcgac gtaaacggcc acaagttcag 4500
cgtgtccggc gagggcgagg gcgatgccac ctacggcaag ctgaccctga agttcatctg 4560
caccaccggc aagctgcccg tgccctggcc caccctcgtg accaccctga cctacggcgt 4620
gcagtgcttc agccgctacc ccgaccacat gaagcagcac gacttcttca agtccgccat 4680
gcccgaaggc tacgtccagg agcgcaccat cttcttcaag gacgacggca actacaagac 4740
ccgcgccgag gtgaagttcg agggcgacac cctggtgaac cgcatcgagc tgaagggcat 4800
cgacttcaag gaggacggca acatcctggg gcacaagctg gagtacaact acaacagcca 4860
caacgtctat atcatggccg acaagcagaa gaacggcatc aaggtgaact tcaagatccg 4920
ccacaacatc gaggacggca gcgtgcagct cgccgaccac taccagcaga acacccccat 4980
cggcgacggc cccgtgctgc tgcccgacaa ccactacctg agcacccagt ccgccctgag 5040
caaagacccc aacgagaagc gcgatcacat ggtcctgctg gagttcgtga ccgccgccgg 5100
gatcactctc ggcatggacg agctgtacaa gtaaagcggc cgcgactcta gatcataatc 5160
agccatacca catttgtaga ggttttactt gctttaaaaa acctcccaca cctccccctg 5220
aacctgaaac ataaaatgaa tgcaattgtt gttgttaaca tcgataaaat aaaagatttt 5280
atttagtctc cagaaaaagg ggggaatgaa agaccccacc tgtaggtttg gcaagctagc 5340
ataacttcgt ataatgtatg ctatacgaag ttattctaga gaaccatcag atgtttccag 5400
ggtgccccaa ggacctgaaa tgaccctgtg ccttatttga actaaccaat cagttcgctt 5460
ctcgcttctg ttcgcgcgct tctgctcccc gagctcaata aaagagccca caacccctca 5520
ctcggggcgc cagtcctccg attgactgag tcgcccgggt acccgtgtat ccaataaacc 5580
ctcttgcagt tgcatccgac ttgtggtctc gctgttcctt gggagggtct cctctgagtg 5640
attgactacc cgtcagcggg ggtctttcat ttgggggctc gtccgggatc gggagacccc 5700
tgcccaggga ccaccgaccc accaccggga ggtaagctgg ctgcctcgcg cgtttcggtg 5760
atgacggtga aaacctctga cacatgcagc tcccggagac ggtcacagct tgtctgtaag 5820
cggatgccgg gagcagacaa gcccgtcagg gcgcgtcagc gggtgttggc gggtgtcggg 5880
gcgcagccat gacccagtca cgtagcgata gcggagtgta tactggctta actatgcggc 5940
atcagagcag attgtactga gagtgcacca tatgcggtgt gaaataccgc acagatgcgt 6000
aaggagaaaa taccgcatca ggcgctcttc cgcttcctcg ctcactgact cgctgcgctc 6060
ggtcgttcgg ctgcggcgag cggtatcagc tcactcaaag gcggtaatac ggttatccac 6120
agaatcaggg gataacgcag gaaagaacat gtgagcaaaa ggccagcaaa aggccaggaa 6180
ccgtaaaaag gccgcgttgc tggcgttttt ccataggctc cgcccccctg acgagcatca 6240
caaaaatcga cgctcaagtc agaggtggcg aaacccgaca ggactataaa gataccaggc 6300
gtttccccct ggaagctccc tcgtgcgctc tcctgttccg accctgccgc ttaccggata 6360
cctgtccgcc tttctccctt cgggaagcgt ggcgctttct catagctcac gctgtaggta 6420
tctcagttcg gtgtaggtcg ttcgctccaa gctgggctgt gtgcacgaac cccccgttca 6480
gcccgaccgc tgcgccttat ccggtaacta tcgtcttgag tccaacccgg taagacacga 6540
cttatcgcca ctggcagcag ccactggtaa caggattagc agagcgaggt atgtaggcgg 6600
tgctacagag ttcttgaagt ggtggcctaa ctacggctac actagaagga cagtatttgg 6660
tatctgcgct ctgctgaagc cagttacctt cggaaaaaga gttggtagct cttgatccgg 6720
caaacaaacc accgctggta gcggtggttt ttttgtttgc aagcagcaga ttacgcgcag 6780
aaaaaaagga tctcaagaag atcctttgat cttttctacg gggtctgacg ctcagtggaa 6840
cgaaaactca cgttaaggga ttttggtcat gagattatca aaaaggatct tcacctagat 6900
ccttttaaat taaaaatgaa gttttaaatc aatctaaagt atatatgagt aaacttggtc 6960
tgacagttac caatgcttaa tcagtgaggc acctatctca gcgatctgtc tatttcgttc 7020
atccatagtt gcctgactcc ccgtcgtgta gataactacg atacgggagg gcttaccatc 7080
```

tggccccagt gctgcaatga taccgcgaga cccacgctca ccggctccag atttatcagc 7140
aataaaccag ccagccggaa gggccgagcg cagaagtggt cctgcaactt tatccgcctc 7200
catccagtct attaattgtt gccgggaagc tagagtaagt agttcgccag ttaatagttt 7260
gcgcaacgtt gttgccattg ctgcaggcat cgtggtgtca cgctcgtcgt ttggtatggc 7320
ttcattcagc tccggttccc aacgatcaag gcgagttaca tgatccccca tgttgtgcaa 7380
aaaagcggtt agctccttcg gtcctccgat cgttgtcaga agtaagttgg ccgcagtgtt 7440
atcactcatg gttatggcag cactgcataa ttctcttact gtcatgccat ccgtaagatg 7500
cttttctgtg actggtgagt actcaaccaa gtcattctga gaatagtgta tgcggcgacc 7560
gagttgctct tgcccggcgt caacacggga taataccgcg ccacatagca gaactttaaa 7620
agtgctcatc attggaaaac gttcttcggg gcgaaaactc tcaaggatct taccgctgtt 7680
gagatccagt tcgatgtaac ccactcgtgc acccaactga tcttcagcat cttttacttt 7740
caccagcgtt tctgggtgag caaaaacagg aaggcaaaat gccgcaaaaa agggaataag 7800
ggcgacacgg aaatgttgaa tactcatact cttccttttt caatattatt gaagcattta 7860
tcagggttat tgtctcatga gcggatacat atttgaatgt atttagaaaa ataaacaaat 7920
aggggttccg cgcacatttc cccgaaaagt gccacctgac gtctaagaaa ccattattat 7980
catgacatta acctataaaa ataggcgtat cacgaggccc tttcgtcttc aagaattcat 8040
accagatcac cgaaaactgt cctccaaatg tgtcccсctc acactcccaa attcgcgggc 8100
ttctgcctct tagaccactc taccctattc cccacactca ccggagccaa agccgcggcc 8160
cttccgtttc tttgcttttg aaagacccca cccgtaggtg gcaagctagc gatgaccctg 8220
ctgattggtt cgctgaccat ttccggggtg cggaacggcg ttaccagaaa ctcagaaggt 8280
tcgtccaacc aaaccgactc tgacggcagt ttacgagaga gatgataggg tctgcttcag 8340
taagccagat gctacacaat taggcttgta catattgtcg ttagaacgcg gctacaatta 8400
atacataacc ttatgtatca tacacatacg atttaggtga cactatagaa tacaagctgg 8460
aagatcttcc agcttgggct gcaggtcgac tctagagtcc gttacataac ttacggtaaa 8520
tggcccgcct ggctgaccgc ccaacgaccc ccgcccattg acgtcaataa tgacgtatgt 8580
tcccatagta acgccaatag ggactttcca ttgacgtcaa tgggtggagt atttacggta 8640
aactgcccac ttggcagtac atcaagtgta tcatatgcca agtacgcccc ctattgacgt 8700
caatgacggt aaatggcccg cctggcatta tgcccagtac atgaccttat gggactttcc 8760
tacttggcag tacatctacg tattagtcat cgctattacc atggtgatgc ggttttggca 8820
gtacatcaat gggcgtggat agcggtttga ctcacgggga tttccaagtc tccaccccat 8880
tgacgtcaat gggagtttgt tttggcacca aaatcaacgg gactttccaa aatgtcgtaa 8940
caactccgcc ccattgacgc aaatgggcgg taggcgtgta cggtgggagg tctatataag 9000
cagagctcgt ttagtgaacc gcgccagtct tccgatagac tgcgtcgccc gggtacccgt 9060
attcccaata aagcctcttg ctgtttgcat ccgaatcgtg gtctcgctgt tccttgggag 9120
ggtctcctct gagtgattga ctacccacga cgggggtctt tcatttgggg gctcgtccgg 9180
gatttggaga cccctgccca gggaccaccg acccaccacc gggaggtaag ctggccagca 9240
acttatctgt gtctgtccga ttgtctagtg tctatgtttg atgttatgcg cctgcgtctg 9300
tactagttag ctaactagct ctgtatctgg cggacccgtg gtggaactga cgagttctga 9360
acacccggcc gcaaccctgg gagacgtccc agggactttg ggggccgttt ttgtggcccg 9420
acctgaggaa gggagtcgat gtggaatccg accccgtcag gatatgtggt tctggtagga 9480
gacgagaacc taaaacagtt cccgcctccg tctgaatttt tgctttcggt ttggaaccga 9540
agccgcgcgt cttgtctgct gcagcgctgc agcatcgttc tgtgttgtct ctgtctgact 9600
gtgtttctgt atttgtctga aaattagggc cagactgtta ccactccctt aagtttgacc 9660
ttaggtcact ggaaagatgt cgagcggatc gctcacaacc agtcggtaga tgtcaagaag 9720
agacgttgg 9729

<210> 57
<211> 7591
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 57
gttaccttct gctctgcaga atggccaacc tttaacgtcg gatggccgcg agacggcacc 60

```
tttaaccgag acctcatcac ccaggttaag atcaaggtct tttcacctgg cccgcatgga    120
cacccagacc aggtccccta catcgtgacc tgggaagcct tggcttttga ccccctccc     180
tgggtcaagc cctttgtaca ccctaagcct ccgcctcctc ttcctccatc cgccccgtct    240
ctccccttg aacctcctcg ttcgaccccg cctcgatcct ccctttatcc agccctcact     300
ccttctctag gcgccggaat tcgttaactc gaggatccac cggtcgccac catggtgagc    360
aagggcgagg agctgttcac cggggtggtg cccatcctgg tcgagctgga cggcgacgta    420
aacggccaca agttcagcgt gtccggcgag ggcgagggcg atgccaccta cggcaagctg    480
accctgaagt tcatctgcac caccggcaag ctgcccgtgc cctggcccac cctcgtgacc    540
accctgacct acggcgtgca gtgcttcagc cgctaccccg accacatgaa gcagcacgac    600
ttcttcaagt ccgccatgcc cgaaggctac gtccaggagc gcaccatctt cttcaaggac    660
gacggcaact acaagacccg cgccgaggtg aagttcgagg gcgacaccct ggtgaaccgc    720
atcgagctga agggcatcga cttcaaggag gacggcaaca tcctggggca caagctggag    780
tacaactaca acagccacaa cgtctatatc atggccgaca agcagaagaa cggcatcaag    840
gtgaacttca agatccgcca caacatcgag gacggcagcg tgcagctcgc cgaccactac    900
cagcagaaca cccccatcgg cgacggcccc gtgctgctgc ccgacaacca ctacctgagc    960
acccagtccg ccctgagcaa agaccccaac gagaagcgcg atcacatggt cctgctggag   1020
ttcgtgaccg ccgccgggat cactctcggc atggacgagc tgtacaagta aagcggccct   1080
aggggtcttt cccctctcgc caaaggaatg caaggtctgt tgaatgtcgt gaaggaagca   1140
gttcctctgg aagcttcttg aagacaaaca acgtctgtag cgaccctttg caggcagcgg   1200
aacccccac ctggcgacag gtgcctctgc ggccaaaagc caccgagttg gttcagctgc    1260
tgcctgaggc tggacgacct cgcggagttc taccggcagt gcaaatccgt cggcatccag   1320
gaaaccagca gcggctatcc gcgcatccat gcccccgaac tgcaggagtg ggaggcacg    1380
atggccgctt tggtcgaggc ggatccggcc attagccata ttattcattg gttatatagc   1440
ataaatcaat attggctatt ggccattgca tacgttgtat ccatatcata atatgtacat   1500
ttatattggc tcatgtccaa cattaccgcc atgttgacat tgattattga ctagttatta   1560
atagtaatca attacggggt cattagttca tagcccatat atggagttcc gcgttacata   1620
acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat tgacgtcaat   1680
aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc aatgggtgga   1740
gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc caagtacgcc   1800
ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt acatgacctt   1860
atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta ccatggtgat   1920
gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg gatttccaag   1980
tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac gggactttcc   2040
aaaatgtcgt aacaactccg ccccattgac gcaaatgggc ggtaggcatg tacggtggga   2100
ggtctatata agcagagctc gtttagtgaa ccgtcagatc gcctggagac gccatccacg   2160
ctgttttgac ctccatagaa gacaccggga ccgatccagc ctccgcggcc ccaagcttac   2220
catgggggca ggtgccaccg gccgcgccat ggacgggccg cgcctgctgc tgttgctgct   2280
tctgggggtg tcccttggag gtgccaagga ggcatgcccc acaggcctgt acacacacag   2340
cggtgagtgc tgcaaagcct gcaacctggg cgagggtgtg gcccagcctt gtggagccaa   2400
ccagaccgtg tgtgagccct gcctggacag cgtgacgttc tccgacgtgg tgagcgcgac   2460
cgagccgtgc aagccgtgca ccgagtgcgt ggggctccag agcatgtcgg cgccgtgcgt   2520
ggaggccgac gacgccgtgt gccgctgcgc ctacggctac taccaggatg agacgactgg   2580
gcgctgcgag gcgtgccgcg tgtgcgaggc gggctcgggc ctcgtgttct cctgccagga   2640
caagcagaac accgtgtgcg aggagtgccc cgacggcacg tattccgacg aggccaacca   2700
cgtggacccg tgcctgccct gcaccgtgtg cgaggacacc gagcgccagc tccgcgagtg   2760
cacacgctgg gccgacgccg agtgcgagga gatccctggc cgttggatta cacggtccac   2820
acccccagag ggctcggaca gcacagcccc cagcacccag gagcctgagg cacctccaga   2880
acaagacctc atagccagca cggtggcagg tgtggtgacc acagtgatgg gcagctccca   2940
gcccgtggtg acccgaggca ccaccgacaa cctcatccct gtctattgct ccatcctggc   3000
tgctgtggtt gtgggccttg tggcctacat agccttcaag aggtggaaca gctgctgagt   3060
cgactctaga ggatccccaa catcgataaa ataaaagatt ttatttagtc tccagaaaaa   3120
ggggggaatg aaagacccca cctgtaggtt tggcaagcta gcttaagtaa cgccattttg   3180
caaggcatgg aaaaatacat aactgagaat agagaagttc agatcaaggt caggaacaga   3240
tggaacagct gaatatgggc caaacaggat atctgtggta agcagttcct gccccggctc   3300
agggccaaga acagatggaa cagctgaata tgggccaaac aggatatctg tggtaagcag   3360
ttcctgcccc ggctcagggc caagaacaga tggtccccag atgcggtcca gccctcagca   3420
gtttctagag aaccatcaga tgtttccagg gtgccccaag gacctgaaat gaccctgtgc   3480
cttatttgaa ctaaccaatc agttcgcttc tcgcttctgt tcgcgcgctt ctgctccccg   3540
```

agctcaataa aagagcccac aacccctcac tcggggcgcc agtcctccga ttgactgagt 3600
cgcccgggta cccgtgtatc caataaaccc tcttgcagtt gcatccgact tgtggtctcg 3660
ctgttccttg ggagggtctc ctctgagtga ttgactaccc gtcagcgggg gtctttcatt 3720
tgggggctcg tccgggatcg ggagacccct gcccagggac caccgaccca ccaccgggag 3780
gtaagctggc tgcctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 3840
cccggagacg gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 3900
cgcgtcagcg ggtgttggcg ggtgtcgggg cgcagccatg acccagtcac gtagcgatag 3960
cggagtgtat actggcttaa ctatgcggca tcagagcaga ttgtactgag agtgcaccat 4020
atgcggtgtg aaataccgca cagatgcgta aggagaaaat accgcatcag gcgctcttcc 4080
gcttcctcgc tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct 4140
cactcaaagg cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg 4200
tgagcaaaag gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc 4260
cataggctcc gcccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga 4320
aacccgacag gactataaag ataccaggcg tttccccctg gaagctccct cgtgcgctct 4380
cctgttccga ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg 4440
gcgctttctc atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag 4500
ctgggctgtg tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat 4560
cgtcttgagt ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac 4620
aggattagca gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac 4680
tacggctaca ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc 4740
ggaaaaagag ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt 4800
tttgtttgca agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc 4860
ttttctacgg ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg 4920
agattatcaa aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca 4980
atctaaagta tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca 5040
cctatctcag cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag 5100
ataactacga tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac 5160
ccacgctcac cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc 5220
agaagtggtc ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct 5280
agagtaagta gttcgccagt taatagtttg cgcaacgttg ttgccattgc tgcaggcatc 5340
gtggtgtcac gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg 5400
cgagttacat gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc 5460
gttgtcagaa gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat 5520
tctcttactg tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag 5580
tcattctgag aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat 5640
aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg 5700
cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca 5760
cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga 5820
aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc 5880
ttcctttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata 5940
tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 6000
ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc 6060
acgaggccct ttcgtctcgc gcgtttcggt gatgacggtg aaaacctctg acacatgcag 6120
ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag 6180
ggcgcgtcag cgggtgttgg cgggtgtcgg ggctggctta actatgcggc atcagagcag 6240
attgtactga gagtgcacca tatggacata ttgtcgttag aacgcggcta caattaatac 6300
ataaccttat gtatcataca catacgattt aggtgacact atagaactcg actctagagt 6360
ccgttacata acttacggta aatggcccgc ctggctgacc gcccaacgac ccccgcccat 6420
tgacgtcaat aatgacgtat gttcccatag taacgccaat agggactttc cattgacgtc 6480
aatgggtgga gtatttacgg taaactgccc acttggcagt acatcaagtg tatcatatgc 6540
caagtacgcc ccctattgac gtcaatgacg gtaaatggcc cgcctggcat tatgcccagt 6600
acatgacctt atgggacttt cctacttggc agtacatcta cgtattagtc atcgctatta 6660
ccatggtgat gcggttttgg cagtacatca atgggcgtgg atagcggttt gactcacggg 6720
gatttccaag tctccacccc attgacgtca atgggagttt gttttggcac caaaatcaac 6780
gggactttcc aaaatgtcgt aacaactccg ccccattgac gcaaatgggc ggtaggcgtg 6840
tacggtggga ggtctatata agcagagctc gtttagtgaa ccgcgccagt cttccgatag 6900
actgcgtcgc ccgggtaccc gtattcccaa taaagcctct tgctgtttgc atccgaatcg 6960
tggtctcgct gttccttggg agggtctcct ctgagtgatt gactacccac gacgggggtc 7020

tttcatttgg gggctcgtcc gggatttgga gacccctgcc cagggaccac cgacccacca 7080
ccgggaggta agctggccag caacttatct gtgtctgtcc gattgtctag tgtctatgtt 7140
tgatgttatg cgcctgcgtc tgtactagtt agctaactag ctctgtatct ggcggacccg 7200
tggtggaact gacgagttct gaacacccgg ccgcaaccct gggagacgtc ccagggactt 7260
tgggggccgt ttttgtggcc cgacctgagg aagggagtcg atgtggaatc cgaccccgtc 7320
aggatatgtg gttctggtag gagacgagaa cctaaaacag ttcccgcctc cgtctgaatt 7380
tttgctttcg gtttggaacc gaagccgcgc gtcttgtctg ctgcagcgct gcagcatcgt 7440
tctgtgttgt ctctgtctga ctgtgtttct gtatttgtct gaaaattagg gccagactgt 7500
taccactccc ttaagtttga ccttaggtca ctggaaagat gtcgagcgga tcgctcacaa 7560
ccagtcggta gatgtcaaga agagacgttg g 7591

<210> 58
<211> 2870
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 58
gctagcataa cttcgtataa tgtatgctat acgaagttat tctagagaac catcagatgt 60
ttccagggtg ccccaaggac ctgaaatgac cctgtgcctt atttgaacta accaatcagt 120
tcgcttctcg cttctgttcg cgcgcttctg ctccccgagc tcaataaaag agcccacaac 180
ccctcactcg gggcgccagt cctccgattg actgagtcgc ccgggtaccc gtgtatccaa 240
taaaccctct tgcagttgca tccgacttgt ggtctcgctg ttccttggga gggtctcctc 300
tgagtgattg actacccgtc agcgggggtc tttcatttgg gggctcgtcc gggatcggga 360
gacccctgcc cagggaccac cgacccacca ccgggaggta agctggctgc ctcgcgcgtt 420
tcggtgatga cggtgaaaac ctctgacaca tgcagctccc ggagacggtc acagcttgtc 480
tgtaagcgga tgccgggagc agacaagccc gtcagggcgc gtcagcgggt gttggcgggt 540
gtcggggcgc agccatgacc cagtcacgta gcgatagcgg agtgtatact ggcttaacta 600
tgcggcatca gagcagattg tactgagagt gcaccatatg cggtgtgaaa taccgcacag 660
atgcgtaagg agaaaatacc gcatcaggcg ctcttccgct tcctcgctca ctgactcgct 720
gcgctcggtc gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt 780
atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 840
caggaaccgt aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 900
gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 960
ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac 1020
cggatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata gctcacgctg 1080
taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 1140
cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag 1200
acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 1260
aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt 1320
atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 1380
atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 1440
gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 1500
gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 1560
ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 1620
ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 1680
tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt 1740
accatctggc cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt 1800
atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc 1860
cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 1920
tagtttgcgc aacgttgttg ccattgctgc aggcatcgtg gtgtcacgct cgtcgtttgg 1980
tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt 2040
gtgcaaaaaa gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc 2100
agtgttatca ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt 2160

aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 2220
gcgaccgagt tgctcttgcc cggcgtcaac acgggataat accgcgccac atagcagaac 2280
tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 2340
gctgttgaga tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt 2400
tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 2460
aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 2520
catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 2580
acaaataggg gttccgcgca catttccccg aaaagtgcca cctgacgtct aagaaaccat 2640
tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtcttcaaga 2700
attcatacca gatcaccgaa aactgtcctc caaatgtgtc cccctcacac tcccaaattc 2760
gcgggcttct gcctcttaga ccactctacc ctattcccca cactcaccgg agccaaagcc 2820
gcggcccttc cgtttctttg cttttgaaag accccacccg taggtggcaa 2870

<210> 59
<211> 3097
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 59
gctagcttaa gtaacgccat tttgcaaggc atggaaaaat acataactga gaatagagaa 60
gttcagatca aggtcaggaa cagatggaac agctgaatat gggccaaaca ggatatctgt 120
ggtaagcagt tcctgccccg gctcagggcc aagaacagat ggaacagctg aatatgggcc 180
aaacaggata tctgtggtaa gcagttcctg ccccggctca gggccaagaa cagatggtcc 240
ccagatgcgg tccagccctc agcagtttct agagaaccat cagatgtttc cagggtgccc 300
caaggacctg aaatgaccct gtgccttatt tgaactaacc aatcagttcg cttctcgctt 360
ctgttcgcgc gcttctgctc cccgagctca ataaaagagc ccacaacccc tcactcgggg 420
cgccagtcct ccgattgact gagtcgcccg ggtacccgtg tatccaataa accctcttgc 480
agttgcatcc gacttgtggt ctcgctgttc cttgggaggg tctcctctga gtgattgact 540
acccgtcagc gggggtcttt catttggggg ctcgtccggg atcgggagac ccctgcccag 600
ggaccaccga cccaccaccg ggaggtaagc tggctgcctc gcgcgtttcg gtgatgacgg 660
tgaaaacctc tgacacatgc agctcccgga gacggtcaca gcttgtctgt aagcggatgc 720
cgggagcaga caagcccgtc agggcgcgtc agcgggtgtt ggcgggtgtc ggggcgcagc 780
catgacccag tcacgtagcg atagcggagt gtatactggc ttaactatgc ggcatcagag 840
cagattgtac tgagagtgca ccatatgcgg tgtgaaatac cgcacagatg cgtaaggaga 900
aaataccgca tcaggcgctc ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt 960
cggctgcggc gagcggtatc agctcactca aaggcggtaa tacggttatc cacagaatca 1020
ggggataacg caggaaagaa catgtgagca aaaggccagc aaaaggccag gaaccgtaaa 1080
aaggccgcgt tgctggcgtt tttccatagg ctccgccccc ctgacgagca tcacaaaaat 1140
cgacgctcaa gtcagaggtg gcgaaacccg acaggactat aaagatacca ggcgtttccc 1200
cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc 1260
gcctttctcc cttcgggaag cgtggcgctt tctcatagct cacgctgtag gtatctcagt 1320
tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg aaccccccgt tcagcccgac 1380
cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc cggtaagaca cgacttatcg 1440
ccactggcag cagccactgg taacaggatt agcagagcga ggtatgtagg cggtgctaca 1500
gagttcttga agtggtggcc taactacggc tacactagaa ggacagtatt tggtatctgc 1560
gctctgctga agccagttac cttcggaaaa agagttggta gctcttgatc cggcaaacaa 1620
accaccgctg gtagcggtgg tttttttgtt tgcaagcagc agattacgcg cagaaaaaaa 1680
ggatctcaag aagatccttt gatcttttct acggggtctg acgctcagtg gaacgaaaac 1740
tcacgttaag ggattttggt catgagatta tcaaaaagga tcttcaccta gatcctttta 1800
aattaaaaat gaagttttaa atcaatctaa agtatatatg agtaaacttg gtctgacagt 1860
taccaatgct taatcagtga ggcacctatc tcagcgatct gtctatttcg ttcatccata 1920
gttgcctgac tccccgtcgt gtagataact acgatacggg agggcttacc atctggcccc 1980
agtgctgcaa tgataccgcg agacccacgc tcaccggctc cagatttatc agcaataaac 2040

cagccagccg gaagggccga gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag 2100
tctattaatt gttgccggga agctagagta agtagttcgc cagttaatag tttgcgcaac 2160
gttgttgcca ttgctgcagg catcgtggtg tcacgctcgt cgtttggtat ggcttcattc 2220
agctccggtt cccaacgatc aaggcgagtt acatgatccc ccatgttgtg caaaaaagcg 2280
gttagctcct tcggtcctcc gatcgttgtc agaagtaagt tggccgcagt gttatcactc 2340
atggttatgg cagcactgca taattctctt actgtcatgc catccgtaag atgcttttct 2400
gtgactggtg agtactcaac caagtcattc tgagaatagt gtatgcggcg accgagttgc 2460
tcttgcccgg cgtcaacacg ggataatacc gcgccacata gcagaacttt aaaagtgctc 2520
atcattggaa aacgttcttc ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc 2580
agttcgatgt aacccactcg tgcacccaac tgatcttcag catcttttac tttcaccagc 2640
gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca 2700
cggaaatgtt gaatactcat actcttcctt tttcaatatt attgaagcat ttatcagggt 2760
tattgtctca tgagcggata catatttgaa tgtatttaga aaaataaaca aataggggtt 2820
ccgcgcacat ttccccgaaa agtgccacct gacgtctaag aaaccattat tatcatgaca 2880
ttaacctata aaaataggcg tatcacgagg ccctttcgtc ttcaagaatt cataccagat 2940
caccgaaaac tgtcctccaa atgtgtcccc ctcacactcc caaattcgcg ggcttctgcc 3000
tcttagacca ctctacccta ttccccacac tcaccggagc caaagccgcg gcccttccgt 3060
ttctttgctt ttgaaagacc ccacccgtag gtggcaa 3097

<210> 60
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 60
ggctagagaa ttccaggtaa gatgggcgat cccctcacct gg 42

<210> 61
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Primer

<400> 61
ttgggtactc ctcgctaggt tc 22

<210> 62
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 62
caggtaag 8

<210> 63
<211> 512
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 63
cgcccgaaca gggacctgag aggggcgcag accctacctg ttgaacctgg ctgatcgtag 60
gatccccggg acagcagagg agaacttaca gaagtcttct ggaggtgttc ctggccagaa 120
cacaggagga caggtaagat gggagaccct ttgacatgga gcaaggcgct caagaagtta 180
gagaaggtga cggtacaagg gtctcagaaa ttaactactg gtaactgtaa ttgggcgcta 240
agtctagtag acttatttca tgataccaac tttgtaaaag aaaaggactg gcagctgagg 300
gatgtcattc cattgctgga agatgtaact cagacgctgt caggacaaga aagagaggcc 360
tttgaaagaa catggtgggc aatttctgct gtaaagatgg gcctccagat taataatgta 420
gtagatggaa aggcatcatt ccagctccta agagcgaaat atgaaaagaa gactgctaat 480
aaaaagcagt ctgagccctc tgaagaatat ct 512

<210> 64
<211> 514
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 64
cgcccgaaca gggacctgag aggggcgcag accctacctg ttgaacctgg ctgatcgtag 60
gatccccggg acagcagagg agaacttaca gaagtcttct ggaggtgttc ctggccagaa 120
cacaggagga caggtaagat tgggagaccc tttgacattg gagcaaggcg ctcaagaagt 180
tagagaaggt gacggtacaa gggtctcaga aattaactac tggtaactgt aattgggcgc 240
taagtctagt agacttattt catgatacca actttgtaaa agaaaaggac tggcagctga 300
gggatgtcat tccattgctg gaagatgtaa ctcagacgct gtcaggacaa gaaagagagg 360
cctttgaaag aacatggtgg gcaatttctg ctgtaaagat gggcctccag attaataatg 420
tagtagatgg aaaggcatca ttccagctcc taagagcgaa atatgaaaag aagactgcta 480
ataaaaagca gtctgagccc tctgaagaat atct 514

<210> 65
<211> 522
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
oligonucleotide

<400> 65
cgcccgaaca gggacctgag aggggcgcag accctacctg ttgaacctgg ctgatcgtag 60
gatccccggg acagcagagg agaacttaca gaagtcttct ggaggtgttc ctggccagaa 120
cacaggagga caggtaagat tgggagaccc tttgacattg gagcaaggcg ctcaagaagt 180
tagagaaggt gacggtacaa gggtctcaga aattaactac tggtaactgt aattgggcgc 240
taagtctagt agacttattt cattgatacc aactttgtaa aagaaaagga ctggcagctg 300
agggattgtc attccattgc tggaagattg taactcagac gctgtcagga caagaaagag 360

```
aggcctttga aagaacattg gtgggcaatt tctgctgtaa agattgggcc tccagattaa    420
taattgtagt agattggaaa ggcatcattc cagctcctaa gagcgaaata ttgaaaagaa    480
gactgctaat aaaaagcagt ctgagccctc tgaagaatat ct                       522
```

# United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division

Application deficiencies found during scanning:

☐ Page(s)__________ of ______________________________ were not present for scanning. (Document title)

☐ Page(s)__________ of ______________________________ were not present for scanning. (Document title)

Sequence Listing are Parts of Specs.

☐ ***Scanned copy is best available.***